Urnas de Pólya y su Conexión con los Árboles Aleatorios

Rémi BESSON

22 de Junio de 2015

English summary

My work deals with Pólya urns and their connections to random trees widely used in computer science in order to structure data or analyse algorithms.

In Pólya urns theory we study urns of colored balls with replacement schemes. We draw randomly balls and replace them adding others according to the scheme. Our main aim is to know the asymptotic distribution of each ball color.

In random trees theory the main issues are to count the number of leaves, the tree size, etc... which give us an idea of the efficiency of an algorithm.

The Pólya urns will help us to get asymptotic distributions of these tree characteristics, they provide an easy, elegant and global way to resolve these problems.

Índice general

1 Agradecimientos

Quiero dar las gracias a mi tutor el Señor Fernando López Blázquez por su paciencia y por sus consejos sobre el programa Latex como sobre la sintaxis española que han hecho mas legible y estético mi trabajo. También me gustaría destacar su disponibilidad y amabilidad a lo largo del curso.

También le agradezco, aunque no lo conozca en persona, al Señor Mahmoud (2009) cuyo libro muy claro y didáctico me ha permitido descubrir nuevas áreas de las matemáticas. Me ha mostrado la importancia de tener un espíritu sintético aprovechando el desarollo de una teoría matemática para resolver problemas en otra.

Finalmente agradezco el apoyo de mi familia y de mis amigos que han hecho posible la realización de este trabajo e inolvidable mi estancia en Sevilla como estudiante Erasmus.

2
Introducción

En este trabajo desarrollaremos la teoría de las urnas de Pólya y sus conexiones con los árboles aleatorios.

Las urnas de Pólya fueron introducidas en los años 20 por los matemáticos George Pólya y Eggenberger con el objetivo de modelar el contagio de una enfermedad. Sin embargo, esas urnas tienen muchas más aplicaciones: en informática, biología, epidemiología, física (las urnas de Ehrenfest pueden ser reducidas a una urna de Polya), etc...
¿Qué son las urnas de Pólya? Una urna de Pólya contiene bolas de k colores distintos. En cada etapa se mezcla bien la urna y se extrae una bola al azar. Si la bola escogida es de color i añadimos $a_{i,1}$ bolas de color 1, $a_{i,2}$ bolas de color 2,..., $a_{i,k}$ bolas de color k. Se repite el procedimiento indefinidamente.

Los $a_{i,j}$ pueden ser aleatorios o determinísticos, negativos o positivos.

Se suele representar el esquema de la urna con una matriz:

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,k} \\ a_{2,1} & a_{2,2} & \dots & a_{2,k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k,1} & a_{k,2} & \dots & a_{k,k} \end{pmatrix}.$$

Primero estudiaremos las condiciones de sostenibilidad de una urna. Es decir cuáles son las hipótesis sobre las condiciones iniciales y sobre la matriz asociada a nuestra urna para que se pueda repetir indefinidemente el procedimiento de extracción de bolas.

Luego desarrollaremos los diversos estudios que han sido presentados a lo largo del siglo XX, del modelo más básico de Pólya y Eggenberger (para una urna dicromática) a los teoremas de Smythe (para una urna con un número cualquiera k de colores distintos de bolas).

Finalmente usaremos esos diversos resultados para estudiar las propiedades de varios tipos de árboles aleatorios. Contaremos las hojas de un árbol binario de búsqueda, las rotaciones de un árbol de franja equilibrada, el tamaño de un árbol m-ario de búsqueda etc...

3
Urnas de Pólya

3.1. Sostenibilidad

Un esquema de urnas se dice que es sostenible si siempre es posible extraer bolas y continuar indefinidamente con la regla de reemplazamiento establecida. La condición de sostenibilidad garantiza que se pueden perpetuar las extracciones indefinidamente y por tanto se trata de una condición necesaria para el estudio de propiedades asintóticas de las urnas de Pólya.

Nota 1. *En el caso particular en que la urna contenga solo dos colores, i.e. $k = 2$, supondremos durante todo el trabajo que tales colores son blanco y azul y denotamos W_0 y B_0 respectivamente al número inicial de bolas blancas y de azules en la urna. En este caso la matriz asociada al esquema de extracciones se denotará* $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

La sostenibilidad de una urna puede depender de las condiciones iniciales. Por ejemplo, la urna cuya matriz es $A = \begin{pmatrix} -1 & -1 \\ 1 & 1 \end{pmatrix}$ no es sostenible si $W_0 \geq B_0$ pues es posible extraer sucesivamente bolas de color blanco hasta que no quede ninguna de color azul, a partir de ese momento no es posible continuar con la regla de reemplazamento establecida pues no es posible extraer ninguna bola azul. Sin embargo, la urna es sostenible si $W_0 < B_0$.

A continuación, vamos a dar caracterizaciones de las matrices sostenibles en dimensión dos. Se puede encontrar tal estudio en el libro de Mahmoud (2009).

1) Está claro que si la matriz tiene cuatro elementos negativos no puede ser sostenible. En efecto, en este caso en cada extracción la urna va teniendo menos bolas y en el momento en que se vacíe de un determinado color no es posible continuar.

2) Si la matriz tiene tres elementos negativos la urna no puede ser sostenible. De hecho, cualquier matriz que tenga dos elementos negativos en la misma columna no puede ser sostenible. Por ejemplo, si la matriz es $\begin{pmatrix} - & + \\ - & + \end{pmatrix}$ siempre quitamos bolas blancas

en cada etapa en el momento que no queden más bolas blancas no se podrá continuar por lo que el modelo no puede ser sostenible.

3) Con dos elementos negativos. Si los dos elementos negativos están en la misma columna la urna no es sostenible tal como se vió anteriormente.

Tampoco es sostenible el caso $\begin{pmatrix} + & - \\ - & + \end{pmatrix}$. En efecto, por ejemplo si solo extraemos bolas blancas, quitamos todas las bolas azules hasta tener un problema.

Ahora vamos a estudiar el caso

$$\begin{pmatrix} - & - \\ + & + \end{pmatrix}.$$

Consideremos el caso más crítico, es decir, el camino que vacía la urna. Primero, W_0 tiene que ser un múltiplo de $|a|$, en efecto, en caso contrario cogeríamos bolas blancas hasta quitar todas las bolas blancas y la urna no sería sostenible. Cogemos $\frac{W_0}{|a|}$ bolas blancas y entonces nuestra urna ya no tiene bolas blancas y tiene $B_0 - W_0\frac{|b|}{|a|}$ bolas azules. Entonces necesitamos que $B_0 - W_0\frac{|b|}{|a|} > 0$ para que la urna sea sostenible. La próxima bola cogida será necesariamente una azul. Entonces tenemos ahora c bolas blancas y $B_0 - W_0\frac{|b|}{|a|} + d$ bolas azules en la urna. Volvemos a vaciar las bolas blancas (es el camino crítico) entonces c tiene que ser un múltiplo de $|a|$ y nos encontramos con una urna que tiene 0 bolas blancas y $B_0 - W_0\frac{|b|}{|a|} + d - c\frac{b}{a}$ bolas azules. De nuevo para que la urna sea sostenible necesitamos $B_0 - W_0\frac{|b|}{|a|} + d - c\frac{b}{a} > 0$. De nuevo vamos a coger una bola azul en la siguiente etapa y la composición de la urna será: c bolas blancas y $B_0 - W_0\frac{|b|}{|a|} + 2d - c\frac{b}{a}$ bolas azules y así sucesivamente.

En el i-ésimo estado crítico tenemos 0 bolas blancas y $B_0 - W_0\frac{|b|}{|a|} + (i-1)d - (i-1)c\frac{b}{a}$ bolas azules. Entonces necesitamos $B_0 - W_0\frac{|b|}{|a|} + (i-1)d - (i-1)c\frac{b}{a} > 0$, $\forall i \geq 1$ es decir, $B_0 - W_0\frac{|b|}{|a|} > (i-1)[c\frac{b}{a} - d]$.

Como i puede ser muy grande y $B_0 - W_0\frac{|b|}{|a|} > 0$ necesitamos $c\frac{b}{a} - d < 0 \Leftrightarrow cb - ad > 0 \Leftrightarrow det(A) = ad - bc < 0$.

En resumen, tenemos las condiciones siguientes para que la urna sea sostenible:

a) W_0 y c son múltiplos de $|a|$
b) $det(A) \leq 0$

c) $det\begin{pmatrix} a & b \\ W_0 & B_0 \end{pmatrix} < 0$

El caso $\begin{pmatrix} + & + \\ - & - \end{pmatrix}$ es simétrico invirtiendo los colores.

Queda el caso

$$\begin{pmatrix} - & + \\ + & - \end{pmatrix}.$$

Vamos a seguir el camino que vacía las bolas blancas (hacemos un razonamiento similar por el camino que vacía las bolas azules). Primero, tenemos que W_0 debe ser un múltiplo de $|a|$, luego vaciamos las bolas blancas de la urna, tenemos 0 bolas blancas y $B_0 + \frac{b}{|a|}W_0$ bolas azules. En la próxima etapa cogeremos necesariamente una bola azul y la urna tiene ahora c blancas y $B_0 + \frac{b}{|a|}W_0 + d$ azules (entonces necesitamos $B_0 + \frac{b}{|a|}W_0 + d > 0$ pero esta condición no aporta nada más porque B_0 es un múltiplo de $|d|$). Volvemos a vaciar las bolas blancas de la urna. Entonces, tenemos la condición que c sea un múltiplo de $|a|$ y nuestra urna tiene ahora 0 blancas y $B_0 + \frac{b}{|a|}W_0 + d + b\frac{c}{|a|}$. Siguiendo así infinitamente en el i-ésimo estado crítico nos encontramos con una urna con 0 blancas y $B_0 + \frac{b}{|a|}W_0 + (i-1)d + (i-1)b\frac{c}{|a|}$ que es siempre positivo.

Entonces las únicas condiciones son:

a) W_0 y c son múltiplos de $|a|$
b) B_0 y b son múltiplos de $|d|$
c) b y c son positivos

4) Ahora con un solo menos. Consideremos los casos simétricos $\begin{pmatrix} + & - \\ + & + \end{pmatrix}$ y $\begin{pmatrix} + & + \\ - & + \end{pmatrix}$. No son sostenibles si cogemos siempre un color de bola hasta vaciar la otra. Entonces para el modelo $\begin{pmatrix} + & - \\ + & + \end{pmatrix}$ necesitamos que $W_0 = 0$ y $c = 0$ (porque nunca tenemos que encontrarnos con blancas en la urna, en caso contrario, podríamos vaciar las azules), y entonces también nos hace falta que $B_0 > 0$.

Ahora consideremos el caso $\begin{pmatrix} - & + \\ + & + \end{pmatrix}$ y está claro que W_0 tiene que ser un múltiplo de a. Cogiendo sólo blancas tenemos una urna con composición 0 blancas y $B_0 + b\frac{W_0}{a}$

azules que tiene que ser estrictamente positivo. Entonces si $b = 0$ necesitamos $B_0 > 0$. En el i-ésimo estado crítico nos encontramos con una urna de composición 0 blancas y $B_0 + b\frac{W_0}{a} + c + b(i-1)\frac{c}{|a|}$ que tiene que ser estrictamente positivo lo cual no añade condiciones.

El caso $\begin{pmatrix} + & + \\ + & - \end{pmatrix}$ es simétrico.

3.2. Urnas de Pólya-Eggenberger

Vamos a presentar el modelo de urnas llamado Pólya-Eggenberger que es una urna de Pólya con dos colores y que tiene $\begin{pmatrix} s & 0 \\ 0 & s \end{pmatrix}$ por matriz (con $s > 0$). Denotaremos $\pi_n = W_n + B_n$ el número de bolas en la urna después de n extracciones, siendo W_n el número de bolas blancas y B_n el número de bolas azules. En este caso tenemos $\pi_n = \pi_0 + sn$. Tenemos el siguiente resultado:

Teorema 3.2.1. *Eggenberger and Pólya (1923). Sea W_n^* el número de bolas blancas seleccionadas en las n extracciones. Tenemos para $0 \leq k \leq n$*

$$P[W_n^* = k] = \frac{W_0(W_0+s)...(W_0+(k-1)s)B_0(B_0+s)...(B_0+(n-k-1)s)}{\pi_0(\pi_0+s)...(\pi_0+(n-1)s)}\binom{n}{k}.$$

Demostración. Si cogemos k blancas en las n extracciones significa que hemos cogido $n-k$ azules. Sea i_j el número de la extracción en la que se escoge la j-ésima blanca, $j = 1, ..., k$. La probabilidad de este camino particular es:

$$\frac{B_0}{\pi_0} \times \frac{B_0+s}{\pi_1} \times \frac{B_0+2s}{\pi_2} \times ... \times \frac{B_0+(i_1-2)s}{\pi_{i_1-2}} \times \frac{W_0}{\pi_{i_1-1}} \times \frac{B_0+(i-1)s}{\pi_{i_1}} \times ... \times \frac{B_0+(i_2-3)s}{\pi_{i_2-2}} \times$$
$$\frac{W_0+s}{\pi_{i_2-1}} \times \frac{B_0+(i_2-2)s}{\pi_{i_2}} \times ... \times \frac{B_0+(n-k-1)s}{\pi_{n-1}}$$

y esta expresión no depende de cuándo se escogen las blancas (es decir, no depende de los índices i_l). Como hay $\binom{n}{k}$ maneras de elegir los momentos donde se cogen las blancas, de aquí el resultado. □

Corolario 3.2.1. *Eggenberger and Pólya (1923). Sea W_n el número de bolas blancas en la urna de Pólya Eggenberger después de n extracciones. Entonces*

$$E[W_n] = \frac{W_0}{\pi_0}sn + W_0,$$

$$Var[W_n] = \frac{W_0 B_0 s^2 n(sn+\pi_0)}{\pi_0^2(\pi_0+s)}.$$

Teorema 3.2.2. *Sea W_n^* el número de bolas blancas seleccionadas en n extracciones. Entonces*

$$\frac{W_n^*}{n} \to_D \beta\left(\frac{W_0}{s}, \frac{B_0}{s}\right).$$

Demostración. Suponemos B_0 y W_0 estrictamente positivos. Tenemos

$$\begin{aligned} P[W_n^* = k] &= \frac{W_0(W_0+s)...(W_0+(k-1)s)B_0(B_0+s)...(B_0+(n-k-1)s)}{\pi_0(\pi_0+s)...(\pi_0+(n-1)s)}\binom{n}{k} \\ &= \frac{\Gamma\left(k+\frac{W_0}{s}\right)\Gamma\left(n-k+\frac{B_0}{s}\right)\Gamma\left(\frac{\pi_0}{s}\right)}{\Gamma\left(\frac{W_0}{s}\right)\Gamma\left(\frac{B_0}{s}\right)\Gamma\left(n+\frac{\pi_0}{s}\right)} \times \frac{n!}{k!(n-k)!}. \end{aligned}$$

Entonces para $x \in [0,1]$ tenemos

$$P(W_n^* \le nx) = \sum_{k=0}^{\lfloor nx \rfloor} \frac{\Gamma\left(k+\frac{W_0}{s}\right)\Gamma\left(n-k+\frac{B_0}{s}\right)\Gamma\left(\frac{\pi_0}{s}\right)}{\Gamma\left(\frac{W_0}{s}\right)\Gamma\left(\frac{B_0}{s}\right)\Gamma\left(n+\frac{\pi_0}{s}\right)} \times \frac{\Gamma(n+1)}{\Gamma(k+1)\Gamma(n-k+1)}.$$

Ahora utilizamos la fórmula de Stirling que nos da una aproximación de la función gamma. En efecto se recuerda la aproximación de Stirling:

$$\Gamma(x+1) = x! \approx \sqrt{2\pi x}(\frac{x}{e})^x$$

con $x \to \infty$. Por lo cual:

$$\begin{aligned} \frac{\Gamma(x+r)}{\Gamma(x+s)} &= \frac{\sqrt{2\pi(x+r-1)}}{\sqrt{2\pi(x+s-1)}}\frac{(\frac{x+r}{e})^{x+r}}{(\frac{x+s}{e})^{x+s}} \\ &= \sqrt{\frac{x+r-1}{x+s-1}}e^{s-r}(x+r)^{x+r}\frac{1}{(x+s)^{x+s}} \\ &= \sqrt{\frac{1+\frac{r-1}{x}}{1+\frac{s-1}{x}}}e^{s-r}\frac{x^{x+r}}{x^{x+s}}(1+\frac{r}{x})^{x+r}\frac{1}{(1+\frac{s}{x})^{x+s}} \\ &\approx x^{r-s} \end{aligned}$$

pasando al límite y recordando que $(1+\frac{r}{x})^{x+r} = e^{(x+r)ln(1+\frac{r}{x})} \approx e^r$

$$P(W_n^* \le nx) = \frac{\Gamma\left(\frac{\pi_0}{s}\right)}{\Gamma\left(\frac{W_0}{s}\right)\Gamma\left(\frac{B_0}{s}\right)} \sum_{k=0}^{\lfloor nx \rfloor} \frac{\Gamma(n+1)}{\Gamma(n+\frac{\pi_0}{s})}\frac{\Gamma(n-k+\frac{B_0}{s})}{\Gamma(n-k+1)}\frac{\Gamma(k+\frac{W_0}{s})}{\Gamma(k+1)}$$

Además utilizando la approximación de Stirling obtenemos:

$$\frac{\Gamma(n+1)}{\Gamma(n+\frac{\pi_0}{s})} \approx n^{1-\frac{\pi_0}{s}};\ \ \frac{\Gamma(n-k+\frac{B_0}{s})}{\Gamma(n-k+1)} \approx (n-k)^{\frac{B_0}{s}-1};\ \ \frac{\Gamma(k+\frac{W_0}{s})}{\Gamma(k+1)} \approx k^{\frac{W_0}{s}-1},$$

con lo cual tomando límite ($n \to \infty$):

$$P\left[\frac{W_n^*}{n} \leq x\right] \to \frac{\Gamma\left(\frac{\pi_0}{s}\right)}{\Gamma\left(\frac{W_0}{s}\right)\Gamma\left(\frac{B_0}{s}\right)} \int_0^x u^{\frac{W_0}{s}-1}(1-u)^{\frac{B_0}{s}-1} du.$$

□

3.3. La urna de Bernard Friedman

El modelo de urna de Pólya ha sido ampliado en los años 40 por Bernstein (1940), Savkevich (1940) y Friedman (1949) con urnas de matriz:

$$\begin{pmatrix} s & a \\ a & s \end{pmatrix}.$$

Éstas se conocen como urnas de Friedman.

Podemos notar que esta matriz es de suma constante sobre las filas, lo cuál permite tener resultados más estéticos. Como la tasa de crecimiento de la urna es constante tenemos que

$$\pi_n = W_n + B_n = \pi_0 + (a+s)n.$$

Supondremos siempre que $s \neq a$ sino la urna es degenerada, no hay nada aleatorio en esta urna y entonces no tiene interés.

Teorema 3.3.1. *Friedman (1949). Sea W_n el número de bolas blancas en una urna de Friedman después de n extracciones. La función generatriz de momentos $\phi_n(t) = E[\exp(W_n t)]$ verifica la ecuación diferencial siguiente:*

$$\phi_{n+1}(t) = e^{at}\left[\phi_n(t) + \frac{e^{(s-a)t}-1}{\pi_n}\phi_n'(t)\right].$$

Demostración. Denotamos 1_n^W y 1_n^B respectivamente a las funciones indicadoras de los eventos coger una blanca en el n-ésimo paso y coger azul en el n-ésimo paso. El número de blancas después de $n+1$ pasos es el número de blancas después de n pasos más el número (negativo o positivo) de bolas blancas que se añaden tras haber cogido la $(n+1)$-ésimo bola o sea:

$$W_{n+1} = W_n + s1_{n+1}^W + a1_{n+1}^B = W_n + (s-a)1_{n+1}^W + a$$

$$E\left[e^{W_{n+1}t} \mid W_n\right] = e^{(W_n+a)t} E\left[e^{(s-a)1_{n+1}^W t} \mid W_n\right]$$

$$\begin{aligned} E\left[e^{(s-a)1_{n+1}^W t} \mid W_n\right] &= E\left[e^{(s-a)1_{n+1}^W t} \mid W_n, 1_n^W = 0\right] P(1_n^W = 0 \mid W_n) \\ &\quad + E\left[e^{(s-a)1_{n+1}^W t} \mid W_n, 1_n^W = 1\right] P(1_n^W = 1 \mid W_n) \\ &= P(1_n^W = 0 \mid W_n) + e^{(s-a)t} P(1_n^W = 1 \mid W_n) \\ &= \left(1 - \frac{W_n}{\pi_n}\right) + \frac{W_n}{\pi_n} e^{(s-a)t}, \end{aligned}$$

tomando esperanza obtenemos:

$$\begin{aligned} \phi_{n+1}(t) &= E\left[e^{W_{n+1}t}\right] \\ &= E\left[E\left[e^{W_{n+1}t} \mid W_n\right]\right] \\ &= E\left[e^{(W_n+a)t}\left(1 - \frac{W_n}{\pi_n}\right) + e^{(W_n+a)t}\frac{W_n}{\pi_n} e^{(s-a)t}\right] \\ &= e^{at} E\left[e^{W_n t}\left(1 - \frac{W_n}{\pi_n} + \frac{W_n}{\pi_n} e^{(s-a)t}\right)\right] \\ &= e^{at}\left(\phi_n(t) + E\left[e^{W_n t} W_n \frac{(e^{(s-a)t} - 1)}{\pi_n}\right]\right) \\ &= e^{at}\left(\phi_n(t) + \frac{e^{(s-a)t} - 1}{\pi_n}\phi_n'(t)\right) \end{aligned}$$

(nótese que π_n no es aleatorio y por lo cual se puede sacar de la esperanza). □

Una ecuación funcional así no es especialmente fácil de resolver para cualquier valor de a y s. Sin embargo, soluciones explícitas existen para $a = 0$ o $s = 0$. Se puede simplificar esta ecuación con el cambio (propuesto por Friedman) siguiente:

$$\chi_n(t) = \left(1 - e^{-t(s-a)}\right)^{\delta+\alpha n} \phi_n(t)$$

con $\delta = \dfrac{\pi_0}{s-a}$ y $\alpha = \dfrac{s+a}{s-a}$.

En efecto, tenemos entonces que

$$\phi_n(t) = \left(1 - e^{-t(s-a)}\right)^{-\delta-\alpha n} \chi_n(t) \tag{3.1}$$

derivando se obtiene que

$$\phi_n'(t) = (-\delta-\alpha n)(s-a)e^{-t(s-a)}\left(1-e^{-t(s-a)}\right)^{-\delta-\alpha n-1}\chi_n(t)+\left(1-e^{-t(s-a)}\right)^{-\delta-\alpha n}\chi_n'(t).$$

Se tiene lo siguiente:

$$\begin{aligned}
\chi_{n+1}(t) &= \left(1-e^{-t(s-a)}\right)^{\delta+\alpha(n+1)}\phi_{n+1}(t)\\
&= \left(1-e^{-t(s-a)}\right)^{\delta+\alpha(n+1)}e^{at}\left[\phi_n(t)+\frac{e^{(s-a)t}-1}{\pi_n}\phi_n'(t)\right]\\
&= \left(1-e^{-t(s-a)}\right)^{\delta+\alpha(n+1)}e^{at}\left[\left(1-e^{-t(s-a)}\right)^{-\delta-\alpha n}\chi_n(t)\right.\\
&+ \frac{e^{(s-a)t}-1}{\pi_n}\left((-\delta-\alpha n)(s-a)e^{-t(s-a)}\left(1-e^{-t(s-a)}\right)^{-\delta-\alpha n-1}\chi_n(t)\right.\\
&+ \left.\left.\left(1-e^{-t(s-a)}\right)^{-\delta-\alpha n}\chi_n'(t)\right)\right]\\
&= \left(1-e^{-t(s-a)}\right)^{\alpha}e^{at}\left[\chi_n(t)\left(1-\frac{e^{t(s-a)}-1}{\pi_n}(\delta+\alpha n)(s-a)e^{-t(s-a)}\left(1-e^{-t(s-a)}\right)^{-1}\right)\right.\\
&+ \left.\frac{e^{(s-a)t}-1}{\pi_n}\chi_n'(t)\right]
\end{aligned}$$

se tiene

$$1-\frac{e^{t(s-a)}-1}{\pi_n}(\delta+\alpha n)(s-a)e^{-t(s-a)}\left(1-e^{-t(s-a)}\right)^{-1} = 1+\frac{-1}{\pi_n}\frac{\pi_0+(s+a)n}{s-a}(s-a) = 0$$

y por otro lado

$$\left(e^{(s-a)t}-1\right)e^{at} = e^{st}-e^{at} = \left(1-e^{-t(s-a)}\right)e^{st}$$

de lo cual obtenemos

$$\chi_{n+1}(t) = \frac{e^{st}}{\pi_n}\left(1-e^{-t(s-a)}\right)^{\alpha+1}\chi_n'(t). \tag{3.2}$$

Veremos aplicaciones de esta ecuación funcional para los árboles recursivos. Ahora vamos a presentar la teoría asintótica de las urnas de Friedman (desarollado por Freedman (1965)).

Teorema 3.3.2. *Freedman (1965). Sea W_n el número de bolas blancas después de n pasos en una urna de Friedman no degenerada y sostenible. Sea $\rho = \frac{s-a}{s+a}$, y suponemos $\rho < \frac{1}{2}$ entonces:*

$$\frac{W_n - \frac{1}{2}(s+a)n}{\sqrt{n}} \to_D N\left(0, \frac{(s-a)^2}{4(1-2\rho)}\right).$$

Si $\rho = \frac{1}{2}$ tenemos

$$\frac{W_n - B_n}{\sqrt{n \log n}} \to_D N(0, (s-a)^2).$$

Si $\rho > \frac{1}{2}$ el comportamiento es completamente diferente:

$$\frac{W_n - B_n}{n^\rho} \to_D \beta\left(\frac{W_0}{s}, \frac{B_0}{s}\right).$$

Nota. Se puede notar que en el caso $\rho \leq \frac{1}{2}$ la distribución asintótica no depende de las condiciones iniciales de la urna. Al revés, en el caso $\rho > \frac{1}{2}$ (la urna de Pólya-Eggenberger es un caso especial con $\rho = 1$) la distribución asintótica depende fuertemente de la composición inicial de la urna. Demostraremos un resultado más general en la sección siguiente.

3.4. La urna de Bagchi-Pal

Para generalizar la urna de Friedman es natural romper la simetría de la matriz correspondiente. Estamos ahora con el caso más general:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Vamos a añadir algunas condiciones. Antes de todo, la urna tiene que ser sostenible para desarollar una teoría asintótica. Suponemos de nuevo que $a + b = c + d = K$ (para tener un crecimiento fijo en cada paso de la urna). También suponemos como garantía de sostenibilidad que $b > 0$, $c > 0$ y si $a < 0$, entonces a divide a W_0 y c; de manera similar si $d < 0$, d divide B_0 y b. Excluimos los casos $b = c = 0$ que corresponde a una urna de Pólya-Eggenberger. Se excluye el caso $a = c$ pues no hay aleatoriedad. También se excluye el caso donde un elemento diagonal es 0. Dicho de otra forma, o ninguno de los elementos de la matriz es nulo o los dos elementos diagonales son nulos.

Proposición 1. *(Bagchi and Pal (1985)). Sea W_n el número de bolas blancas después de n pasos. Entonces*

$$E[W_n] = \frac{c}{b+c} Kn + o(n). \tag{3.3}$$

Si $a - c < \frac{1}{2}K$,

$$V[W_n] = \frac{bcK(a-c)^2}{(b+c)^2(K-2(a-c))}n + o(n), \tag{3.4}$$

y si $a - c = \frac{1}{2}K$,

$$V[W_n] = \frac{bc}{K}n\log n + O(1).$$

Demostración. Conociendo el número de bolas blancas después de n pasos sabemos que la urna después de $n+1$ pasos sólo puede tener $W_n + a$ (en el caso que cogemos una blanca) o $W_n + c$ (en el caso de que cogemos una azul) bolas blancas. Por eso tenemos las igualdades:

$$P(W_{n+1} = W_n + a \mid W_n) = \frac{W_n}{\pi_n} \tag{3.5}$$

y

$$P(W_{n+1} = W_n + c \mid W_n) = 1 - \frac{W_n}{\pi_n} \tag{3.6}$$

lo cual nos da:

$$\begin{aligned} E[W_{n+1} \mid W_n] &= (W_n + a)P[W_{n+1} = W_n + a \mid W_n] + (W_n + c)P[W_{n+1} = W_n + c \mid W_n] \\ &= (W_n + a)\frac{W_n}{\pi_n} + (W_n + c)\left(1 - \frac{W_n}{\pi_n}\right) \\ &= \left(1 + \frac{a-c}{\pi_n}\right)W_n + c. \end{aligned}$$

Tomando esperanza

$$E[E[W_{n+1} \mid W_n]] = E[W_{n+1}] = (1 + \frac{a-c}{\pi_n})E[W_n] + c$$

lo cual es casi una sucesión geométrica por eso introducimos el cambio

$$Y_n = W_n - \frac{c}{b+c}\pi_n. \tag{3.7}$$

y obtenemos la relación

$$\begin{aligned} E[Y_{n+1}] &= E[W_{n+1}] - \frac{c}{b+c}\pi_{n+1} \\ &= \left(1 + \frac{a-c}{\pi_n}\right)E[W_n] + c - \frac{c}{b+c}\pi_{n+1} \\ &= \left(1 + \frac{a-c}{\pi_n}\right)\left(E[Y_n] + \frac{c}{b+c}\pi_n\right) + c - \frac{c}{b+c}\pi_{n+1} \\ &= \left(1 + \frac{a-c}{\pi_n}\right)E[Y_n] \end{aligned}$$

porque

$$\begin{aligned}\left(1+\frac{a-c}{\pi_n}\right)\left(\frac{c}{b+c}\pi_n\right)+c-\frac{c}{b+c}\pi_{n+1}=0\\ \Leftrightarrow \frac{c}{b+c}\pi_n+\frac{(a-c)c}{b+c}+c-\frac{c}{b+c}\pi_{n+1}=0\\ \Leftrightarrow c\pi_n+(a-c)c+c(b+c)-c\pi_{n+1}=0\\ \Leftrightarrow \pi_n+a+b-\pi_{n+1}=0\end{aligned}$$

lo cual es cierto utilizando la hipótesis de crecimiento fijo de la urna de Bagchi y Pal.

Utilizando esta relación obtenemos

$$\begin{aligned}E[Y_n]&=\left(1+\frac{a-c}{\pi_{n-1}}\right)E[Y_{n-1}]\\ &=\left(1+\frac{a-c}{\pi_{n-1}}\right)\left(1+\frac{a-c}{\pi_{n-2}}\right)E[Y_{n-2}]\\ &=\left(W_0-\frac{c}{b+c}\pi_0\right)\prod_{j=0}^{n-1}\left(1+\frac{a-c}{\pi_j}\right)\\ &=\left(W_0-\frac{c}{b+c}\pi_0\right)\prod_{j=0}^{n-1}\left(1+\frac{a-c}{\pi_0+jK}\right)\\ &=\left(W_0-\frac{c}{b+c}\pi_0\right)\prod_{j=0}^{n-1}\frac{j+\dfrac{\pi_0+a-c}{K}}{j+\dfrac{\pi_0}{K}}\\ &=\left(W_0-\frac{c}{b+c}\pi_0\right)\prod_{j=1}^{n}\frac{j+\dfrac{\pi_0+a-c}{K}-1}{j+\frac{\pi_0}{K}-1}\\ &=\left(W_0-\frac{c}{b+c}\pi_0\right)\frac{\left(\frac{\pi_0}{K}-1\right)!\left(n+\frac{\pi_0+a-c}{K}-1\right)!}{\left(\dfrac{\pi_0+a-c}{K}-1\right)!\left(n+\dfrac{\pi_0}{K}-1\right)!)}\\ &=\left(W_0-\frac{c}{b+c}\pi_0\right)\frac{\Gamma\left(\frac{\pi_0}{K}\right)\Gamma\left(n+\dfrac{\pi_0+a-c}{K}\right)}{\Gamma\left(\dfrac{\pi_0+a-c}{K}\right)\Gamma\left(n+\dfrac{\pi_0}{K}\right)}\end{aligned}$$

Utilizamos la approximación de Stirling y obtenemos:

$$\frac{\Gamma(n+\frac{\pi_0+a-c}{K})}{\Gamma(n+\frac{\pi_0}{K})}=n^{\frac{a-c}{K}}+O(n^{\frac{a-c}{K}-1}),$$

con lo cual:

$$E[Y_n] = O(n^{\frac{a-c}{K}})$$

y como $a - c < K$, $E[Y_n] = O(n)$ y por tanto

$$E[W_n] = \frac{c}{b+c}\pi_n + E[Y_n] = \frac{cK}{b+c}n + O(n).$$

No desarrollaremos completamente el cálculo de la varianza. Utilizando (3.5) y (3.6):

$$P(W_{n+1}^2 = (W_n + a)^2 \mid W_n) = \frac{W_n}{\pi_n},$$
$$P(W_{n+1}^2 = (W_n + c)^2 \mid W_n) = 1 - \frac{W_n}{\pi_n}.$$

Utilizando (3.7) obtenemos la recurrencia:

$$E[Y_{n+1}^2] = \left(1 + \frac{2(a-c)}{\pi_n}\right) E[Y_n^2] + \frac{(b-c)(a-c)^2}{(b+c)\pi_n} E[Y_n] + \frac{bc(a-c)^2}{(b+c)^2}.$$

Si $a - c < \frac{1}{2}K$, la recurrencia asintóticamente tiene una solución lineal. Si $a - c = \frac{1}{2}K$ esta recurrencia se simplifica como

$$E[Y_{n+1}^2] = \frac{\pi_{n+1}}{\pi_n} E[Y_n^2] + \frac{b^2 - c^2}{\pi_n} E[Y_n] + bc$$

que tiene una solución particular superlineal.

□

Corolario 3.4.1. *En una urna de Bagchi y Pal no degenerada, si $a - c < \frac{1}{2}K$ tenemos*

$$\frac{W_n}{n} \to_P \frac{cK}{b+c}.$$

Demostración. En la desigualdad de Chebychev:

$$P(\mid W_n - E[W_n] \mid > \epsilon) \leq \frac{Var[W_n]}{\epsilon^2}, \forall \epsilon > 0,$$

cambiamos ϵ por $\epsilon E[W_n]$ y entonces obtenemos:

$$P\left(\mid \frac{W_n}{E[W_n]} - 1 \mid > \epsilon\right) \leq \frac{Var[W_n]}{\epsilon^2 E[W_n]^2} \to 0$$

si $n \to +\infty$ por (3.4) y (3.3).

Entonces

$$\frac{W_n}{E[W_n]} \to_P 1,$$

además (3.3) nos da

$$\frac{E[W_n]}{n} \to \frac{cK}{b+c}$$

y usando estas dos últimas relaciones obtenemos:

$$\frac{W_n}{n} \to_P \frac{cK}{b+c}.$$

□

Vamos a presentar un resultado que nos será útil en el estudio de los árboles binarios de búsqueda.

Teorema 3.4.1. *(Bagchi and Pal (1985)) Sea W_n el número de bolas blancas en una urna de Bagchi y Pal después de n pasos, donde se añade K bolas a cada paso. Si $a - c < \frac{1}{2}K$ tenemos*

$$W_n^* := \frac{W_n - \frac{cK}{b+c}n}{\sqrt{n}} \to_D N\left(0, \frac{bcK(a-c)^2}{(b+c)^2(K-2(a-c))}\right).$$

Si $a - c = \frac{1}{2}K$

$$W_n^* := \frac{W_n - \frac{cK}{b+c}n}{\sqrt{nln(n)}} \to_D N\left(0, \frac{bc}{K}\right).$$

3.5. Urnas ampliadas y Teoremas de Smythe

Hasta ahora nos habíamos quedado en el caso de una urna con bolas de dos colores. Es natural generalizar nuestra teoría a urnas k colores. Consideremos a partir de ahora una urna con bolas de k colores y sea $A = [a_{i,j}]_{1\leq i,j\leq k}$ su matriz correspondiente con las mismas hipótesis que la urna de Bagchi y Pal, es decir, que se permiten elementos diagonales negativos sólo si la urna es sostenible, y se supone además que $\sum_{j=1}^{k} a_{i,j} = K$, $\forall i \in \{1, 2, ..., k\}$.

Denotamos $X_n^{(i)}$ el número de bolas de color i después de n pasos.
Si $a_{j,j} < 0$ tenemos que $a_{j,j} | X_0^{(j)}$ como consecuencia de la sostenibilidad.

Una urna de matriz $A = [a_{i,j}]_{1\leq i,j\leq k}$ es una urna ampliada si y solo si verifica las condiciones:

1) La urna es sostenible.

2) $\forall i \in \{1, \dots k\}$, $\sum_{j=1}^{k} a_{i,j} = \lambda_1$. Llamamos λ_1 al valor propio principal de A (es el valor propio de mayor parte real)

3) La matriz correspondiente sólo tiene un valor propio real positivo (que es λ_1)

4) Para cada valor propio no principal de A, es decir, los λ_i (con $i = 2, ..., k$), tenemos $Re(\lambda_i) < \frac{1}{2}\lambda_1$.

5) Todos los valores propios son simples.

6) No se puede tener dos valores propios complejos distintos con misma parte real, a no ser que sean conjugados.

7) Los vectores propios son linealmente independientes.

8) No hay valores propios puramente imaginarios.

9) Las componentes del vector propio asociado al vector propio principal (llamado vector propio principal) son positivos.

Teorema 3.5.1. *Smythe (1996). Sea una urna ampliada de k colores, con valor propio principal $\lambda_1 > 0$ y el vector propio asociado normalizado (o sea de norma uno con la norma que suma los valores absolutos de los coeficientes del vector) $v = (v_1, v_2, ..., v_k)$. Sea $X_n^{(i)}$ el número de bolas de color i después de n pasos. Entonces:*

$$\forall i \in \{1, \dots k\}, \ \frac{X_n^{(i)}}{n} \to_P \lambda_1 v_i.$$

Teorema 3.5.2. *Smythe (1996). Suponemos que A es la matriz de una urna ampliada sostenible que tiene por valor propio principal λ_1 y como vector propio izquierdo normalizado asociado V_1^T. Sea $X_n^{(i)}$ el número de bolas de color i en la urna después de n pasos y $X_n = (X_n^{(1)}, ..., X_n^{(k)})^T$ entonces*

$$\frac{1}{\sqrt{n}}(X_n - \lambda_1 n V_1) \to_D N_k(0, \Sigma)$$

donde Σ es una matrix de covarianza límite que no tiene una expresión explícita.

Vamos ahora a hacer la demostración de esos dos teoremas en el caso de una urna ampliada con dos tipos de colores que supondremos al principio por comodidad de suma constante sobre las filas. Consideremos

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

la matriz asociada a una tal urna.

Podemos notar que ya habíamos conseguido resultados de normalidad para la urna de Bagchi y Pal. Vamos a desarrollar aquí una alternativa debida a Smythe (1996), el mayor interés es que se podrá generalizar más facilmente el resultado a $k > 2$.

Podemos notar que las valores propios son $\lambda_1 = K$ y $\lambda_2 = a - c$ utilizando que

$$\begin{cases} \lambda_1 + \lambda_2 & = Tr(A) = a + d \\ \lambda_1 \times \lambda_2 & = det(A) = ad - bc \end{cases}$$

El caso $\lambda_2 = 0$ es trivial ya que implica que $a = c$ y que el esquema

$$\begin{pmatrix} a & K-a \\ a & K-a \end{pmatrix}$$

no tiene aleatoriedad. Con lo cual supondremos que $a \neq c$.

Sea W_n el número de bolas blancas después de n pasos y B_n el número de bolas azules después de n pasos, $\pi_n = W_n + B_n$ el número total de bolas en la urna después de n pasos. En nuestro caso de crecimiento fijo de la urna tenemos $\pi_n = \pi_0 + Kn$ de lo cual obtenemos

$$\frac{\pi_n}{n} \to_P \lambda_1. \tag{3.8}$$

Denotamos $v = (v_1, v_2)$ y $u = (u_1, u_2)$ los vectores propios izquierdos (vectores filas) respectivos de λ_1 y λ_2.
Sea

$$Z_n = \begin{pmatrix} W_n \\ B_n \end{pmatrix}$$

y

$$X_n = uZ_n.$$

Tenemos que $|X_n|$ es linealmente acotado para n grande :

$$\begin{aligned} |X_n| &= |u_1 W_n + u_2 B_n| \\ &\leq |u_1| W_n + |u_2| B_n \\ &\leq max(|u_1|, |u_2|)(W_n + B_n) \\ &= max(|u_1|, |u_2|)\pi_n \\ &= max(|u_1|, |u_2|)(\pi_0 + \lambda_1 n) \\ &\leq max(|u_1|, |u_2|)(\lambda_1 n + \lambda_1 n) \\ &= 2max(|u_1|, |u_2|)\lambda_1 n \\ &= C_1 n \end{aligned} \tag{3.9}$$

Sea $\mathcal{F}_j$ la sigma algebra generada por W_j. Entonces:

$$\begin{aligned}
E[\nabla X_n \mid \mathcal{F}_{n-1}] &= E[X_n - X_{n-1} \mid \mathcal{F}_{n-1}] \\
&= E[X_n \mid \mathcal{F}_{n-1}] - X_{n-1} \\
&= E[uZ_n \mid \mathcal{F}_{n-1}] - X_{n-1} \\
&= E[u_1 W_n + u_2 B_n \mid \mathcal{F}_{n-1}] - X_{n-1} \\
&= u_1 \left(W_{n-1} + a\frac{W_{n-1}}{\pi_{n-1}} + c\frac{B_{n-1}}{\pi_{n-1}} \right) + u_2 \left(B_{n-1} + b\frac{W_{n-1}}{\pi_{n-1}} + d\frac{B_{n-1}}{\pi_{n-1}} \right) \\
&= (u_1 W_{n-1} + u_2 B_{n-1}) + \frac{1}{\pi_{n-1}} u A^T Z_{n-1} - X_{n-1} \\
&= X_{n-1} + \frac{\lambda_2}{\pi_{n-1}} u Z_{n-1} - X_{n-1} \\
&= \frac{\lambda_2}{\pi_{n-1}} X_{n-1}
\end{aligned}$$

con lo cual $E\left[\nabla X_n - \frac{\lambda_2}{\pi_{n-1}} X_{n-1} \mid \mathcal{F}_{n-1}\right] = 0$ y los $M_n = \nabla X_n - \frac{\lambda_2}{\pi_{n-1}} X_{n-1}$ son diferencias de martingalas.

Utilizando el teorema de la doble esperanza obtenemos que $E[M_n] = 0$.

Vamos ahora a buscar coeficientes $\beta_{j,n}$ tal que

$$V_n = \sum_{j=1}^{n} \beta_{j,n} M_j$$

sea una buena aproximación de X_n. Es decir,

$$V_n = \sum_{j=1}^{n} \beta_{j,n} M_j = X_n + \epsilon_n,$$

donde ϵ_n es el eror pequeño de tal forma que no influye asíntoticamente. Entonces un resultado asíntotico para V_n también vale para X_n. Deducimos que $E[V_n] = 0$

Para determinar la expresión de los $\beta_{j,n}$ desarollamos la igualdad anterior:

$$\begin{aligned}
X_n + \epsilon_n &= V_n \\
&= \beta_{n,n} \left(X_n - X_{n-1} - \frac{\lambda_2}{\pi_{n-1}} X_{n-1} \right) + \beta_{n-1,n} \left(X_{n-1} - X_{n-2} - \frac{\lambda_2}{\pi_{n-2}} X_{n-2} \right) \\
&+ ... + \beta_{1,n} \left(X_1 - X_0 - \frac{\lambda_2}{\pi_0} X_0 \right)
\end{aligned}$$

Tomamos $\beta_{n,n} = 1$ y anulamos el coeficiente de X_{n-1} del cual sacamos

$$\beta_{n-1,n} = 1 + \frac{\lambda_2}{\pi_{n-1}}$$

seguimos anulando el coeficiente de X_{n-2} del cual obtenemos

$$\beta_{n-2,n} = \left(1 + \frac{\lambda_2}{\pi_{n-2}}\right)\left(1 + \frac{\lambda_2}{\pi_{n-1}}\right)$$

Trás calculos que no desarrolaremos y que se pueden encontrar en Mahmoud (2009) obtenemos que

$$\beta_{j,n} = \left(\frac{n}{j}\right)^{\frac{\lambda_2}{\lambda_1}} + O\left(n^{\frac{\lambda_2}{\lambda_1}-1}\right)$$

y

$$\epsilon_n = O\left(n^{\frac{\lambda_2}{\lambda_1}}\right)$$

Lema 3.5.1. *Si* $\lambda_2 < \frac{1}{2}\lambda_1$ *entonces*

$$\frac{X_n}{n} \to_P 0.$$

Demostración. Calculamos la varianza

$$\begin{aligned} Var[V_n] &= E[V_n^2] \\ &= E\left[\sum_{j=1}^{n} \beta_{j,n} M_j\right] \\ &= E\left[\sum_{j=1}^{n} \beta_{j,n}^2 M_j^2\right] + 2E\left[\sum_{1\le r<s\le n} \beta_{r,n}\beta_{s,n} M_r M_s\right] \end{aligned}$$

El segundo término es nulo en efecto $E[\beta_{r,n}\beta_{s,n}M_rM_s \mid F_r] = \beta_{r,n}\beta_{s,n}M_rE[M_s \mid F_r] = 0$ y utilizamos el teorema de la doble esperanza. Entonces

$$\begin{aligned} Var[V_n] &= \sum_{j=1}^{n} \beta_{j,n}^2 E[M_j^2] \\ &= \sum_{j=1}^{n} \beta_{j,n}^2 E\left[\left((X_j - X_{j-1}) - \frac{\lambda_2 X_{j-1}}{\pi_{j-1}}\right)^2\right] \\ &\le \sum_{j=1}^{n} \beta_{j,n}^2 E\left[\left(|X_j - X_{j-1}| + \frac{\lambda_2 |X_{j-1}|}{\pi_{j-1}}\right)^2\right] \\ &\le \sum_{j=1}^{n} \beta_{j,n}^2 E\left[\left(|u(Z_n - Z_{n-1})| + \frac{\lambda_2 C_1 (j-1)}{\lambda_1 (j-1) + \pi_0}\right)^2\right] \end{aligned}$$

utilizando 3.9 en la última desigualdad. Y cada componente de la diferencia $|Z_n - Z_{n-1}| = (|\nabla W_n|; |\nabla B_n|)$ esta acotado por máx$(|a|, |b|, |c|, |d|)$. Con lo cual existe constantes C_2 y C_3 tales que

$$\begin{aligned} Var[V_n] &\leq C_2 \sum_{j=1}^{n} \beta_{j,n}^2 \\ &= C_2 \sum_{j=1}^{n} \left(\left(\frac{n}{j} \right)^{\frac{\lambda_2}{\lambda_1}} + O\left(n^{\frac{\lambda_2}{\lambda_1} - 1} \right) \right) \\ &= C_2 \left[\sum_{j=1}^{n} \left(\frac{n}{j} \right)^{\frac{2\lambda_2}{\lambda_1}} + O\left(n^{\frac{2\lambda_2}{\lambda_1} - 2} \right) + O\left(\frac{n^{\frac{2\lambda_2}{\lambda_1} - 1}}{j^{\frac{\lambda_2}{\lambda_1}}} \right) \right] \\ &\leq C_3 n \end{aligned}$$

utilizando la condición $2\lambda_2 < \lambda_1$. Ahora con la desigualdad de Chebychev obtenemos:

$$P[\mid V_n - E[V_n] \mid > n\epsilon] = P\left[\mid \frac{V_n}{n} \mid > \epsilon \right] \leq \frac{C_3 n}{n^2 \epsilon^2} \to_{n \to \infty} 0$$

o sea $\dfrac{V_n}{n} \to_P 0$. De lo cual deducimos

$$\frac{X_n}{n} = \frac{V_n - \epsilon_n}{n} \to_P 0$$

porque habíamos demostrado ya que $\epsilon_n = O(n^{\frac{\lambda_2}{\lambda_1}}) = O(n)$. □

Teorema 3.5.3. *Smythe (1996) Sea W_n y B_n respectivamente el número de bolas blancas y azules en una urna ampliada dicromática, con valor propio λ_1 y el vector propio izquierdo asociado (v_1, v_2). Seguimos denotando $u = (u_1, u_2)$ vector propio izquierdo de λ_2. Entonces,*

$$\frac{W_n}{n} \to_P \lambda_1 v_1 \tag{3.10}$$

$$\frac{B_n}{n} \to_P \lambda_1 v_2.$$

Demostración. Para cada $y = (y_1, y_2)$ existe α_1 y α_2 tal que $y = \alpha_1(1,1) + \alpha_2 u$. Entonces, teniendo en cuenta que los vectores propios v e u son ortogonales (pues λ_1 y λ_2 se suponen distintos)

$$yv = \alpha_1(1,1)v + \alpha_2 uv = \alpha_1(v_1 + v_2) + 0 = \alpha_1,$$

además habíamos supuesto v de norma uno lo cual nos da la última igualdad. Por lo tanto:

$$\begin{aligned}\frac{1}{n}yZ_n &= \frac{1}{n}\alpha_1(1,1)Z_n + \frac{\alpha_2}{n}uZ_n \\ &= \frac{1}{n}yv(1,1)(W_n,B_n)^T + \alpha_2\frac{X_n}{n} \\ &= \frac{1}{n}yv(W_n+B_n)^T + \alpha_2\frac{X_n}{n}.\end{aligned}$$

Además ya habíamos demostrado que $\dfrac{W_n+B_n}{n} = \dfrac{\pi_n}{n} = \dfrac{\lambda_1 n+\pi_0}{n} \to \lambda_1$ y con el lema anterior tenemos $\dfrac{X_n}{n} \to_P 0$ o sea $\dfrac{1}{n}yZ_n \to_P \lambda_1 yv$.

No tenemos ninguna restrición sobre y, así que escogiendo $y = (1,0)$ obtenemos $\dfrac{W_n}{n} \to_P \lambda_1 v_1$ y cogiendo $y = (0,1)$ obtenemos $\dfrac{B_n}{n} \to_P \lambda_1 v_2$. □

Nota 2. *Habíamos demostrado el Teorema 3.5.1 en dimensión $k = 2$. No desarrollaremos el caso $k > 2$. El interés principal de la demostración que acabamos de presentar es ver la importancia de la hipotesis $\lambda_2 < \dfrac{1}{2}\lambda_1$. Esta hipotesis será a menudo la más complicada de verificar cuando queramos utilizar el teorema de Smythe en el estudio de los árboles aleatorios.*

Ahora vamos a demostrar el teorema 3.5.2 en dimensión dos. Tenemos ya una ley débil de los grandes números para $\dfrac{W_n}{n}$. Hemos de encontrar ahora un teorema central del límite para las urnas ampliadas con crecimiento fijo.
Se recuerda algunas notaciónes:$X_n = u_1W_n + u_2B_n$, donde (u_1,u_2) es el vector propio de λ_2.

Además (v_1,v_2) es el vector propio asociado al valor propio principal λ_1.

Sea

$$W_n^* = W_n - v_1\pi_n.$$

La idea es de aproximar W_n^* por una suma de martingalas para las cuales tenemos un teorema central del límite (ver apéndice).

Sea $V_n^* = \displaystyle\sum_{j=1}^{n} \beta_{j,n}^* M_j^* = W_n^* - \epsilon_n^*$ donde M_j^* son martingalas y ϵ_n^* el error pequeño.

Elegimos los $\beta_{j,n}^*$ tal que V_n^* sea una buena aproximación de W_n^*.

Nótese que W_n es linealmente acotado. En efecto, existe una constante C_4 tal que:

$$W_n \leq \pi_n \leq C_4 n.$$

Con lo cual utilizando 3.10 y 3.8 tenemos

$$\frac{W_n^*}{n} = \frac{W_n}{n} - v_1 \frac{\pi_n}{n} \to_P 0$$

Sea 1_n^W el indicador que vale uno si se escoge una blanca en la n-ésima extracción y 0 en caso contrario. Entonces:

$$\begin{aligned} E\left[W_n \mid \mathcal{F}_{n-1}\right] &= W_{n-1} + aP\left[1_n^W = 1 \mid \mathcal{F}_{n-1}\right] + cP\left[1_n^W = 0 \mid \mathcal{F}_{n-1}\right] \\ &= W_{n-1} + a\frac{W_{n-1}}{\pi_{n-1}} + c\frac{B_{n-1}}{\pi_{n-1}} \\ &= W_{n-1} + a\frac{W_{n-1}}{\pi_{n-1}} + c\frac{\pi_{n-1} - W_{n-1}}{\pi_{n-1}} \end{aligned}$$

Tenemos la relación

$$E[W_n^* - W_{n-1}^* \mid \mathcal{F}_{n-1}] = W_{n-1}^* + v_1\pi_{n-1} + a\frac{W_{n-1}^* + v_1\pi_{n-1}}{\pi_{n-1}}$$

reorganizada como

$$\begin{aligned} E\left[W_n^* - W_{n-1}^* \mid F_{n-1}\right] &= v_1(\pi_{n-1} - \pi_n) + (a-c)\frac{W_{n-1}^*}{\pi_{n-1}} + (a-c)v_1 + c \\ &= -\lambda_1 v_1 + (a-c)\frac{W_{n-1}^*}{\pi_{n-1}} + (a-c)v_1 + c \\ &= \lambda_2 \frac{W_{n-1}^*}{\pi_{n-1}} \end{aligned}$$

recordando que $\lambda_1 = K$, $\lambda_2 = a - c$, $v_1 = \dfrac{c}{K + c - a}$, $v_2 = \dfrac{K - a}{K + c - a}$.

Entonces

$$M_n^* = \nabla W_n^* - \frac{\lambda_2}{\pi_{n-1}} W_{n-1}^*$$

es una diferencia de martingalas.

Los M_n^* son uniformemente acotadas:

$$\begin{aligned} |M_n^*| &\leq |\nabla W_n^*| + \frac{|\lambda_2|}{\pi_{n-1}}|W_{n-1}^*| \\ &= |\nabla W_n - v_1 \nabla \pi_n| + \frac{|\lambda_2|}{\pi_{n-1}}|W_{n-1} - v_1 \pi_{n-1}| \\ &\leq |\nabla W_n| + v_1 \nabla \pi_n + \frac{|\lambda_2|}{\pi_{n-1}}(W_{n-1} + v_1 \pi_{n-1}) \\ &\leq max(|a|, |c|) + \lambda_1 v_1 + |\lambda_2|(1 + v_1) = C_5 \end{aligned} \tag{3.11}$$

También nos hace falta el momento de orden 2 de esas diferencias de martingalas, trás un cálculo tedioso obtenemos:

$$\begin{aligned} E\left[(M_n^*)^2 \mid \mathcal{F}_{n-1}\right] = &\left(a^2 \frac{W_{n-1}}{\pi_{n-1}} + c^2 \frac{B_{n-1}}{\pi_{n-1}}\right) \\ &- \left(2\lambda_1 v_1 + 2\frac{\lambda_2 W_{n-1}^*}{\pi_{n-1}}\right)\left(a\frac{W_{n-1}}{\pi_{n-1}} + c\frac{B_{n-1}}{\pi_{n-1}}\right) \\ &+ \lambda_1^2 v_1^2 + \frac{\lambda_2^2 (W_{n-1}^*)^2}{\pi_{n-1}^2} + 2\frac{\lambda_1 \lambda_2 v_1 W_{n-1}^*}{\pi_{n-1}} \end{aligned}$$

Utilizando (3.8) y (3.5) obtenemos que:
$E\left[(M_n^*)^2 \mid \mathcal{F}_{n-1}\right] \to_P a v_1(a - 2\lambda_1 v_1) + c v_2(c - 2\lambda_1 v_1) + \lambda_1^2 v_1^2 = C_6$

Como ha sido ya hecho antes cálcumos los $\beta_{j,n}^*$ y obtenemos:

$$\beta_{j,n}^* = \prod_{k=j}^{n-1}\left(1 + \frac{\lambda_2}{\pi_k}\right)$$

o sea como en el caso anterior

$$\beta_{j,n}^* = \left(\frac{n}{j}\right)^{\frac{\lambda_2}{\lambda_1}} + O\left(n^{\frac{\lambda_2}{\lambda_1} - 1}\right),$$

además

$$\epsilon_n^* = O\left(n^{\frac{\lambda_2}{\lambda_1}}\right).$$

Vamos ahora a comprobar las hipótesis del teorema central del límite de martingalas (ver apéndice).

Utilizando 3.5 tenemos:

$$\frac{1}{n}\sum_{j=1}^{n} E\left[(\beta_{j,n}^* M_j^*)^2 \mid \mathcal{F}_{j-1}\right] = \frac{1}{n}\sum_{j=1}^{n}(\beta_{j,n}^*)^2 E\left[(M_j^*)^2 \mid \mathcal{F}_{j-1}\right]$$
$$\approx_P \frac{1}{n}\sum_{j=1}^{n}\left(\frac{n}{j}\right)^{2\frac{\lambda_2}{\lambda_1}} C_6$$
$$\to_P \frac{C_6}{1-2\frac{\lambda_2}{\lambda_1}} = \sigma^2$$

Aparece de nuevo aquí la importancia de la hipótesis $\lambda_2 < \frac{1}{2}\lambda_1$ sin la cual no tendríamos convergencia.

Entonces,

$$\frac{1}{n}\sum_{j=1}^{n} E\left[(\beta_{j,n}^* M_j^*)^2 \mid \mathcal{F}_{n-1}\right] \to_P \sigma^2.$$

por lo cual V_n^* satisface la condición de varianza σ^2-condicionada.

Además utilizando (3.11) obtenemos:
$max_{1\le j\le n}\frac{E\left[(\beta_{j,n}^* M_j^*)^2 \mid \mathcal{F}_{j-1}\right]}{n} \le max_{1\le j\le n}\frac{C_5^2 E\left[(\beta_{j,n}^*)^2 \mid \mathcal{F}_{j-1}\right]}{n}$

Recordamos que

$$\beta_{j,n}^* = \left(\frac{n}{j}\right)^{\frac{\lambda_2}{\lambda_1}} + O\left(n^{\frac{\lambda_2}{\lambda_1}-1}\right).$$

Entonces:

$$\max_{1\le j\le n}\frac{E\left[(\beta_{j,n}^* M_j^*)^2 \mid \mathcal{F}_{j-1}\right]}{n} \to_P 0$$

utilizando la hipótesis $\lambda_2 < \frac{1}{2}\lambda_1$, lo cual es una condición equivalente a la condición de Lindeberg (ver Hall and Heyde (1980)).

Utilizando el teorema central del límite para las martingalas, la suma $\frac{V_n^*}{n} = \sum_{j=1}^{n}\beta_{j,n}^*\frac{M_j^*}{\sqrt{n}}$ converge en distribución a una variable aleatoria con función característica $e^{-\sigma^2 t^2/2}$. Es decir:

$$\frac{V_n^*}{\sqrt{n}} = \frac{W_n^* - \epsilon_n^*}{\sqrt{n}} = \frac{W_n - v_1\pi_n - \epsilon_n^*}{\sqrt{n}} \to_D N(0,\sigma^2).$$

Además

$$\frac{-\epsilon_n^*}{\sqrt{n}} = \frac{O\left(n^{\frac{\lambda_2}{\lambda_1}}\right)}{\sqrt{n}} \to_{c.s} 0$$

utilizando de nuevo la hipótesis $\frac{\lambda_2}{\lambda_1} < \frac{1}{2}$. Combinando esos dos últimos resultados:

$$\frac{W_n - v_1\pi_n}{\sqrt{n}} \to_D N(0, \sigma^2).$$

Por fin recordando que

$$\frac{\pi_n}{n} \to \lambda_1$$

obtenemos el resultado buscado:

$$\frac{W_n - \lambda_1 v_1 n}{\sqrt{n}} \to_D N(0, \sigma^2)$$

donde

$$\sigma^2 = \frac{av_1(a - 2\lambda_1 v_1) + cv_2(c - 2\lambda_1 v_1) + \lambda_1^2 v_1^2}{1 - 2\frac{\lambda_2}{\lambda_1}}.$$

□

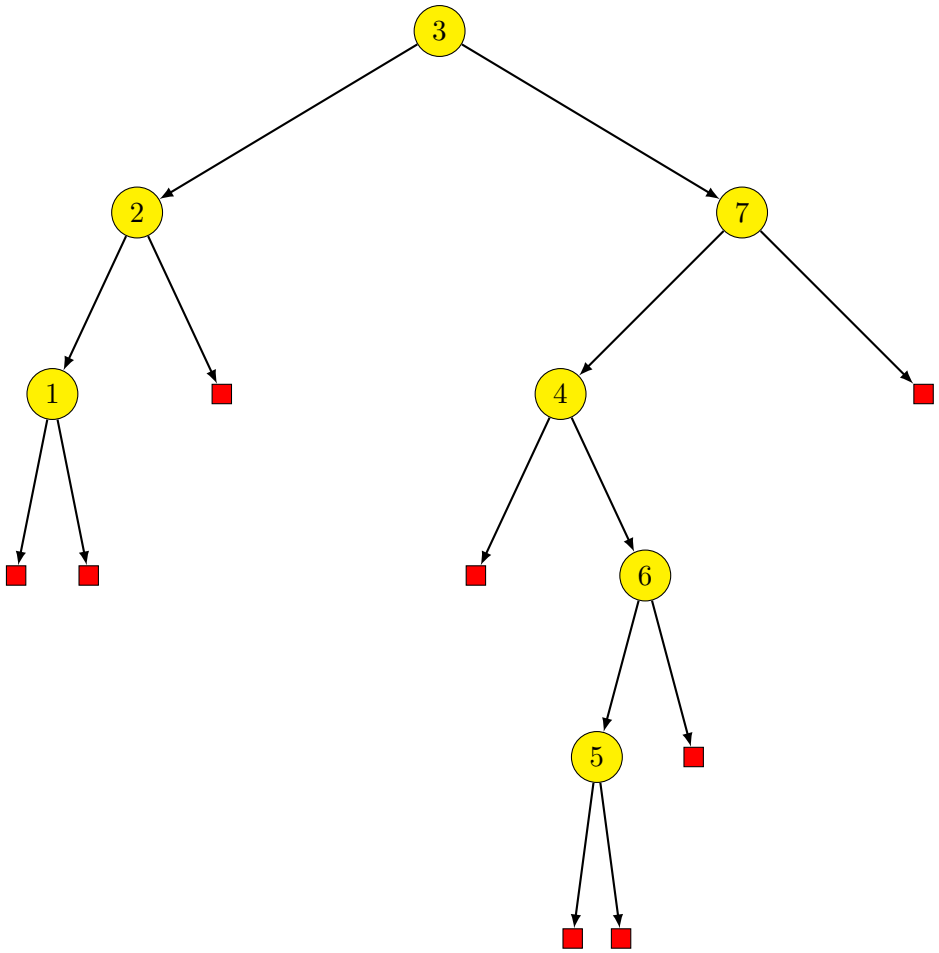
4

Aplicaciones de las urnas de Pólya en informática: los árboles aleatorios

4.1. Los árboles binarios de búsqueda

Vamos a ver varios tipos de árboles aleatorios donde la utilización de la teoría de las urnas de Pólya proporciona resultados interesantes. Consideremos ahora los árboles binarios de búsqueda que son importantes para el almacenamiento de datos o la modelización de algoritmos. Organizar los datos según árbol binario de búsqueda también nos permite localizar más rápidamente el dato que buscamos.

Un árbol binario de búsqueda es una estructura compuesta de nodos cada uno teniendo cero hijos, o un hijo a la izquierda, o un hijo a la derecha, o dos hijos (uno a la derecha y otro a la izquierda). A cada permutación le asociamos un árbol binario de búsqueda con la regla de que cada número de un nodo debe ser superior a los números de los nodos de su izquierda y inferior a los de su derecha. Por ejemplo, sea la permutación $(3, 7, 4, 2, 6, 1, 5)$ el árbol correspondiente es:



En efecto, empezamos con el 3 y luego como el 7 es más grande será su hijo a la derecha, y seguimos construyendo así el árbol. Por ejemplo, si la permutación fuera $(3, 7, 4, 2, 6, 1, 5, 8)$ pondríamos el 8 como hijo derecha del 7. Llamamos nodos internos los círculos amarillos que contienen ya un número. Se extiende el árbol con lo que llamaremos nodos externos (que aparecen en rojo en la figura) de tal forma que cada nodo interno tenga dos hijos (internos o externos). Además llamaremos hojas a los nodos internos que tienen dos hijos externos. Dos nodos serán hermanos si tienen el mismo padre.

La primera pregunta que nos podemos hacer es si la correspondencia que asigna a cada permutación un árbol binario de búsqueda es biyectiva. La respuesta es no. Es obvio que a cada permutación corresponde un sólo árbol binario de búsqueda, pero el recíproco no es cierto. Por ejemplo las permutaciones $\{2, 3, 1, 5, 4, 6\}$ y $\{2, 1, 3, 5, 4, 6\}$ son distintas pero tienen el mismo árbol binario de búsqueda. Es más, si consideramos las permutaciones de los elementos $\{1, ..., n\}$ sabemos que hay $n!$ permutaciones posibles pero solamente hay $(n+1)^{-1}\binom{2n}{n} < n!$ árboles binarios de búsqueda correspondiente, ver Stanley (1999)

En realidad las hojas son nodos de despilfarro, que no se usan. Por eso, nos interesa

conocer el número de hojas para tener una medida de la eficiencia de un árbol binario de búsqueda.

Teorema 4.1.1. *Devroye (1991) Sea L_n el número de hojas en un árbol binario de tamaño n. Entonces*

$$\frac{L_n - \frac{1}{3}n}{\sqrt{n}} \to_D N(0, \frac{2}{45}).$$

Demostración. Coloreemos cada nodo externo que tenga un hermano interno en azul y cada nodo externo que tenga un hermano externo en blanco.
Si la nueva inserción se hace sobre un nodo interno blanco, su hermano se vuelve azul y los dos hijos externos del nodo añadido son blancos.

Ahora, si la nueva inserción se hace sobre un nodo interno azul, se vuelve nodo interno con dos nodos externos blancos.

Lo cual, equivale al esquema:

$$\begin{pmatrix} 0 & 1 \\ 2 & -1 \end{pmatrix}.$$

En efecto, si cogemos una blanca de la urna quitamos dos blancas y añadimos otras dos (es decir, añadimos 0 blancas) y si cogemos una azul quitamos una azul y añadimos dos blancas.

Tenemos una urna de Bagchi y Pal. En efecto, se tiene $a = 0$, $b = 1 > 0$, $c = 2 > 0$ y $d = -1$ por lo cual $a+b = c+d = 1 = K$ y d divide B_0 y b. Además se verifica $a-c < \frac{1}{2}K$ porque $-2 < \frac{1}{2}$.

Denotando W_n el número de bolas blancas en la urna después de n pasos tenemos por Teorema 3.4.1 que

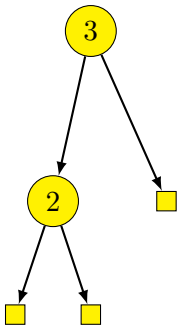
$$\frac{W_n - \frac{c}{b+c}Kn}{\sqrt{n}} = \frac{W_n - \frac{2}{3}n}{\sqrt{n}} \to_D N\left(0, \frac{bcK(a-c)^2}{(b+c)^2(K-2(a-c))}\right) = N\left(0, \frac{8}{45}\right).$$

Finalmente utilizamos que $W_n = 2L_n$ y de aquí el resultado. □

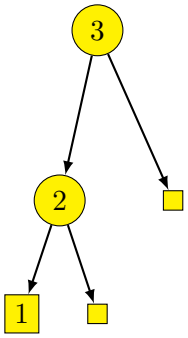
4.2. Árbol de franja equilibrada

Los árboles binarios son bastante eficaces pero se puede aumentar aún más la velocidad de recuperación de datos buscando una forma más compacta de nuestro árbol corrigiendo un desequilibrio. La idea es de operar una rotación de los nodos cuando tenemos un desequilibrio sin romper nuestra regla que los números más grandes están a la derecha de los más pequeños.

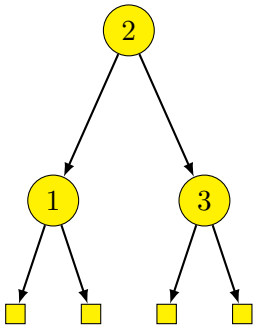
Por ejemplo, consideremos la permutación simple siguiente $(3, 2, 1)$ el árbol correspondiente se construye de la siguiente forma:



y después:

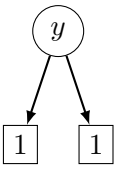


pero tras una rotación obtenemos el árbol siguiente:

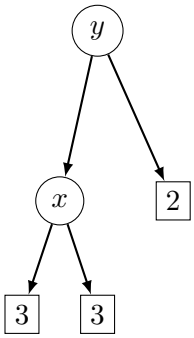


Ampliamos el árbol con nodos externos. Coloreemos en color 1 los nodos externos que tienen en su nivel solamente nodos externos (este color representa generalmente una situación de equilibrio), de color 2 los nodos externos que sólo tienen un hermano interno y los hijos de este nodo interno se colorean en color 3. Tenemos tres casos:

- 1) La nueva entrada (denotado por x) se inserta en un nodo de color 1. En este caso el árbol inicial

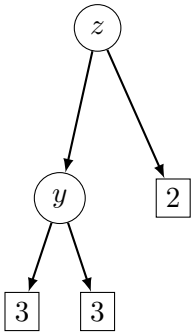


se transforma en

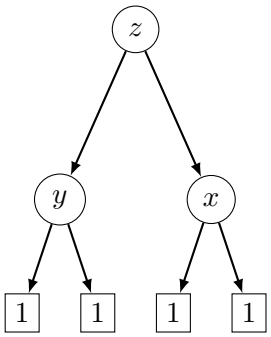


Es decir, hemos retirado dos nodos de color 1 y hemos añadido uno de color 2 y dos de color 3.

- 2) La nueva entrada se inserta en un nodo externo de color 2. En este caso el árbol:

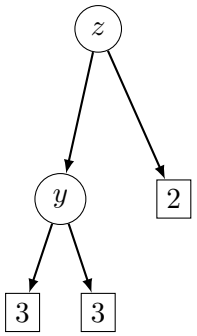


se transforma en

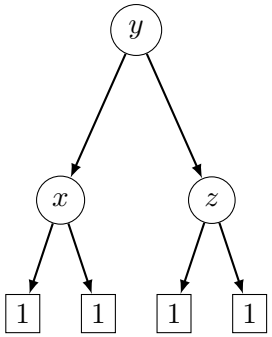


Es decir, hemos añadido cuatro nodos de color 1 y hemos borrado un nodo de color 2 y dos de color 3.

- 3) Si la nueva entrada se inserta en un nodo externo de color 3 se efectua una rotación que equilibra el árbol. El árbol:



se transforma en



Es decir, añadimos cuatro nodos de color 1, y borramos uno de color 2 y dos de color 3. La urna correspondiente a este esquema es:

$$A = \begin{pmatrix} -2 & 1 & 2 \\ 4 & -1 & -2 \\ 4 & -1 & -2 \end{pmatrix}$$

Teorema 4.2.1. *(Mahmoud (1998), Panholzer and Prodinger (1998)) Denotamos R_n el número de rotaciones después de n inserciones en un árbol de franja equilibrada inicialmente vacío. Entonces*

$$\frac{R_n - \frac{2}{7}n}{\sqrt{n}} \to_D N(0, \frac{66}{637})$$

Demostración. La matriz

$$A = \begin{pmatrix} -2 & 1 & 2 \\ 4 & -1 & -2 \\ 4 & -1 & -2 \end{pmatrix}$$

es la de una urna ampliada, en efecto tenemos $\sum_{j=1}^{3} a_{i,j} = 1 = \lambda_1$.

Se trata de un esquema de urna sostenible. Aunque mirando solamente la matriz, podríamos pensar que hace falta hipótesis suplementarias (porque hay números negativos en la matriz). Sin embargo, si nos fijámos en cómo se colorean los nodos externos, nos damos cuenta que nunca se quitan más bolas de color i que el número de bolas de color i que tenemos efectivamente en la urna. Además los valores propios son $\lambda_1 = 1$, $\lambda_2 = 0$ y $\lambda_3 = -6$ por lo cual se cumple la condición $Re(\lambda_2) = 0 < \frac{1}{2}Re(\lambda_1) = \frac{1}{2}$.

Ahora calculamos el vector propio izquierdo asociado al valor propio 1. Es decir buscamos u tal que $u^T A = u^T$. Obtenemos $u = (4, 1, 2)^T$ y lo normalizamos para obtener $v = \frac{1}{7}(4, 1, 2)^T$.

Utilizando el teorema 3.5.1 obtenemos

$$\frac{R_n}{n} \to_P v_3 = \frac{2}{7}$$

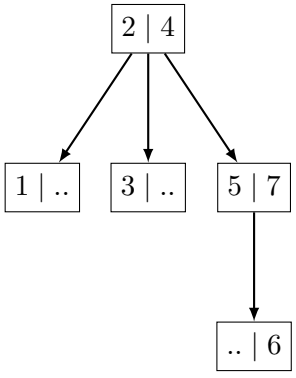
y utilizando el teorema 3.5.2 obtenemos

$$\frac{R_n - \frac{2}{7}n}{\sqrt{n}} \to_D N(0, \sigma^2).$$

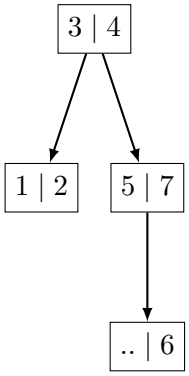
□

4.3. Árbol m-ario de búsqueda

La velocidad de búsqueda se reduce con el número de nodos porque implica más subárboles. Por eso, vamos ahora a considerar árboles m-arios de búsqueda, es decir, que cada nodo puede tener m ramas y que cada nodo tiene $m - 1$ llaves (es decir, $m - 1$ sitios donde se puede insertar un número).
El árbol se construye con la misma regla que anteriormente (los elementos más grandes siempre están a la derecha de los más pequeños). La permutación $(2, 4, 1, 3, 5, 7, 6)$ tiene por árbol ternario ($m = 3$)



Podemos notar que la permutación $(3, 4, 1, 2, 5, 7, 6)$ tiene el mismo número de llaves pero tiene por árbol:



que es más pequeño que el árbol de la permutación $(2, 4, 1, 3, 5, 6, 7)$. Es decir, que el tamaño del árbol m-ario es aleatorio (para $m > 2$) mientras que el tamaño del árbol binario de búsqueda es fijo.

Vamos a ver como podemos aplicar la teoría de las urnas de Pólya a este tipo de árbol. En esta parte me he apoyado sobre el libro de Mahmoud (2009).

Llamamos nodos internos a los nodos que tienen $m - 1$ números (que están ya completos).

A los nodos que no son nodos internos se les llama hojas.

Llamaremos brechas a las posiciones de inserción de una hoja, donde se puede poner un nuevo número. Podemos modelizar un árbol m-ario de búsqueda por una urna de $m-1$ colores. Por eso vamos a colorear las brechas de las hojas. Una hoja con $i - 1$ números tiene i brechas.

Una brecha es de color i si pertenece a una hoja teniendo $i - 1$ números con $i = 1, ..., m-1$. Nótese que si un nodo interno (un nodo que tiene $m-1$ números) tiene menos de m subárboles, le añadimos subárboles vacíos hasta llegar a las m conexiones. Es decir, si tenemos un nodo interno con i hojas reales, le añadimos $m-i$ hojas artificales las cuales corresponden a $m - i$ bolas de color 1 en la urna.

Por ejemplo, nuestro $3-$árbol de la permutación $(2, 4, 1, 3, 5, 6, 7)$ tiene una urna correspondiente compuesta de seis bolas de color 2 (tres hojas con un solo número y cada una tiene dos brechas) y dos bolas de color 1 (los dos hijos del nodo $(5 \mid 6)$ que se añaden) mientras que por la permutación $(3, 4, 1, 2, 5, 6, 7)$ tenemos dos bolas de color 2 y seis bolas de color 1 (porque el nodo interno $(1 \mid 2)$ tiene 3 subárboles que son tres hojas con cada una teniendo una brecha. Además se añade dos subárboles vacíos hijos del nodo $(5 \mid 6)$ y por fin, un hijo árbol vacío del nodo $(3 \mid 4)$)

Un número, insertándose en una hoja de $i - 1 < m - 2$ números, se va añadir a la hoja que tendrá entonces i números y $i + 1$ brechas. Lo cual significa que la urna va a perder i bolas de color i y va a ganar $i + 1$ bolas de color $i + 1$. Esta inserción sólo afecta

las bolas de color i y las de color $i+1$ y no a las otras.

Es un poco diferente si insertamos un número en una hoja que contiene ya $m-2$ números y entonces $m-1$ brechas. En efecto, en este caso, la urna tiene $m-1$ bolas de color $m-1$. La inserción cae en la hoja que tiene ahora $m-1$ números (entonces se vuelve un nodo interno) y por lo cual, está completa y da la luz a m subárboles lo cual equivale en la urna a m bolas de color 1. Es decir, que la urna pierde $m-1$ bolas de color $m-1$ y gana m bolas de color 1. De aquí que la matriz correspondiente sea

$$A = \begin{pmatrix} -1 & 2 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & -2 & 3 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & -3 & 4 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & & & \ddots & & \vdots & \vdots \\ 0 & 0 & \dots & \dots & & & -(m-2) & m-1 \\ m & 0 & \dots & \dots & & & 0 & -(m-1) \end{pmatrix}.$$

El polinomio característico de esta matriz viene dado por (tomando la matriz transpuesta):

$$\chi(\lambda) = \left| \begin{pmatrix} -1-\lambda & 0 & 0 & 0 & 0 & \dots & 0 & m \\ 2 & -2-\lambda & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 3 & -3-\lambda & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & & & \ddots & & \vdots & \vdots \\ 0 & 0 & \dots & \dots & & & -(m-2)-\lambda & 0 \\ 0 & 0 & \dots & \dots & & & m-1 & -(m-1)-\lambda \end{pmatrix} \right|$$
$$= (-1-\lambda)(-2-\lambda)(-3-\lambda)...(-(m-1)-\lambda) + m(-1)^m 2 \times 3 \times ... \times (m-1)$$

desarollando por la primera fila. Y entonces los valores propios verifican

$$(-1)^{m-1}(\lambda+1)(\lambda+2)...(\lambda+m-1) + (-1)^m m! = 0$$
$$\Leftrightarrow (\lambda+1)(\lambda+2)...(\lambda+m-1) = m!$$

Sean λ_1,λ_2,..., λ_{m-1} los valores propios de esta matriz ordenados según su parte real tal que:

$Re(\lambda_1) \geq Re(\lambda_2) \geq ... \geq Re(\lambda_{m-1})$.

Las raices del polínomio mínimo verifican las propiedades siguientes:

1) El valor propio principal es $\lambda_1 = 1$

2) Para m impar, el polínomio mínimo tiene una raiz real negativa, $\lambda_{m-1} = -m-1$.

3) Aparte del valor propio principal y del posible valor propio negativo, todos los otros valores propios son complejos de parte imaginaria no nula.

4) Cada valor propio imaginario tiene su conjugado entre los valores propios.

5) Los valores propios tienen parte real distintas o son conjugados.

6) $Re(\lambda_2) < \frac{1}{2}$ para $m \leq 26$ y $Re(\lambda_2) > \frac{1}{2}$ para $m > 26$.

Primero podemos ver directamente con el polinomio mínimo que 1 es valor propio. Además si m es impar poniendo $\lambda = -m - 1$ en el polinomio mínimo obtenemos

$$(-m) \times (-m+1) \times ... \times (-2) = (-1)^{m-1} m! = m!$$

por lo cual $-m-1$ es valor propio negativo.

Además no hay otro número real, que no sea 1 o $(-m-1)$, que verifique:

$$(\lambda+1)(\lambda+2)...(\lambda+m-1) = m!$$

Entonces como el polinomio mínimo es de coeficientes reales y $\mathbb{C}$ es algebraicamente cerrado, todos los otros valores propios son complejos con parte imaginaria no negativa (lo cual nos da (3)) y sabemos que las raices complejas del polinomio vienen conjugadas (4).

Suponemos que $a + ic$ es valor propio, entonces si $c' > c$ tenemos que

$$\mid a + ic' + 1 \mid\mid a + ic' + 2 \mid ... \mid a + ic' + m - 1 \mid > m!$$

y entonces $a + ic'$ no puede ser valor propio lo cual prueba (5).

Además salvo 1 todos los valores propios tienen parte real menor que 1 por lo cual 1 es valor propio principal. En efecto, suponemos que λ_j es un valor propio complejo (de parte imaginaria no nula) tal que $Re(\lambda_j) \geq 1$ luego:

$$\mid (\lambda_j+1)(\lambda_j+2)...(\lambda_j+m-1) \mid > Re(\lambda_j+1)Re(\lambda_j+2)...Re(\lambda_j+m-1) \geq m!$$

y λ_j no es valor propio, lo cual llega a una contradición entonces $Re(\lambda_j) < 1$.

Admitiremos (6) que necesita bastante más trabajo. Se puede encontrar la demostración en Hosam M. Mahmoud (1995), viene también en este articulo la demostración de los puntos (1) a (5) que acabamos de demostrar.

Para poder aplicar el teorema 3.5.1 vamos a tener que estudiar el vector propio de A asociado al valor propio principal: $\lambda_1 = 1$ (Nótese que el vector propio izquierdo de A asociado a λ_1 es el mismo que el vector propio derecho de A^T asociado a λ_1) .

$$A^T v = \lambda_1 v = v$$
$$\Leftrightarrow -v_1 + m v_{m-1} = v_1; 2v_1 - 2v_2 = v_2; ...; (m-1)v_{m-2} - (m-1)v_{m-1} = v_{m-2}$$

lo cual da:

$$v_{m-1} = \frac{2}{m}v_1; \;\; v_2 = \frac{2}{3}v_1; \;\; v_3 = \frac{3}{4}v_2 = \frac{3}{4} \times \frac{2}{3}v_1 = \frac{2}{4}v_1;$$
$$\dots$$
$$v_{m-1} = \frac{m-2}{m-1}v_{m-2} = \frac{m-2}{m-1}\frac{m-3}{m-2}\dots\frac{2}{3}v_1 = \frac{2}{m-1}v_1,$$

entonces

$$v = (\frac{2}{2}, \frac{2}{3}, ..., \frac{2}{m})$$

La condición de normalizar el vector v se traduce como:

$$v_1 + \frac{2}{3}v_1 + ... + \frac{2}{m}v_1 = 1$$

por lo cual

$$v_1 = \frac{1}{1 + 2(\frac{1}{3} + \frac{1}{4} + ... + \frac{1}{m})} = \frac{1}{2(H_m - 1)}$$

donde H_m es el m-ésimo número harmónico. Luego

$$v_i = \frac{2}{i+1}\frac{1}{2(H_m - 1)} = \frac{1}{(i+1)(H_m - 1)}.$$

Sea $X_n^{(i)}$ el número de bolas de color i después de n pasos. Consideremos el caso $3 \leq m \leq 26$ donde la urna es una urna ampliada. Ahora utilizando el Teorema 3.5.1 de Smythe. Tenemos que:

$$\frac{X_n^{(i)}}{n} \to_P \lambda_1 v_i = \frac{1}{(i+1)(H_m - 1)}.$$

Recordemos que llamamos nodos internos a los nodos que tienen $m-1$ números (es decir, están ya completos). Notaremos I_n el número de nodos internos después de n pasos.

También habíamos definido las hojas como los nodos que no son internos. Denotaremos L_n el número de hojas después de n pasos.

Por fin, se debe notar que hay también lo que llamaremos hojas artificiales, que no están dibujadas porque no llevan números pero que añadimos para que cada nodo interno tenga m ramificaciones. Denotaremos L_n^* el número de hojas artificiales después de n pasos, tenemos $L_n^* = X_n^{(1)}$. Denotamos también $L'_n = L_n + L_n^*$ el número total de hojas.

Si consideremos las hojas de nuestro árbol extendido (las hojas reales como las artificiales) como nodos externos obtenemos la igualdad siguiente: número total de nodos= $I_n + L'_n$ Llamaremos arcos a las lineas que unen los nodos entre ellos. Tenemos que:

1) el número de arcos= mI_n (cada nodo interno tiene m ramificaciones)

2) el número de arcos=número total de nodos$-1 = I_n + L'_n - 1$

Conbinando 1) y 2) obtenemos

$$L'_n = (m-1)I_n + 1$$

.

Lo que nos interesa es el tamaño del árbol, es decir, el número real de nodos (no se cuentan las hojas artificiales):

$$\begin{aligned} S_n &= I_n + L_n \\ &= \frac{L'_n - 1}{m-1} + L_n \\ &= \frac{L'_n - 1}{m-1} + L'_n - X_n^{(1)} \end{aligned}$$

y observando que tenemos $L'_n = \sum_{i=1}^{m-1} \frac{X_n^{(i)}}{i}$ obtenemos

$$S_n = \frac{m}{m-1} \sum_{i=1}^{m-1} \frac{X_n^{(i)}}{i} - X_n^{(1)} - \frac{1}{m-1}$$

con lo cual pasando al límite obtenemos

$$\begin{aligned} \frac{S_n}{n} \to_P & \frac{m}{m-1} \sum_{i=1}^{m-1} \left(\frac{1}{i(i+1)(H_m - 1)} - \frac{1}{2(H_m - 1)} \right) \\ &= \frac{m}{m-1} \frac{1}{H_m - 1} \left(\sum_{i=1}^{m-1} \frac{1}{i} - \sum_{i=1}^{m-1} \frac{1}{i+1} \right) \\ &= \frac{m}{m-1} \frac{1}{H_m - 1} \left(H_m - \frac{1}{m} - H_m + 1 \right) \\ &= \frac{1}{2(H_m - 1)}. \end{aligned}$$

La normalidad asintótica del tamaño del árbol se establece en el siguiente resultado.

Teorema 4.3.1. *(Chern and Hwang (2001)). Sea S_n el tamaño de un árbol m-ario de búsqueda de una permutación de $(1, ..., n)$ y suponemos $3 \leq m \leq 26$ entonces*

$$\frac{S_n - \frac{n}{2(H_m - 1)}}{\sqrt{n}} \to_D N(0, \sigma_m^2).$$

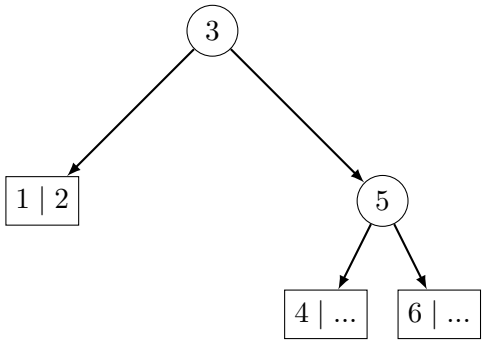
La varianza σ_m^2 puede ser obtenida por recurrencia. H.Mahmoud (1992) da una fórmula exacta. Si $m > 26$ la urna no es una urna extendida en el sentido de Smythe (1996).

4.4. Árboles aleatorios binarios de páginas

Otra generalización de los árboles binarios de búsqueda son los árboles binarios de páginas en las cuales las hojas son cubos que pueden contener hasta $c = 2b$ números. Cada nodo interno tiene un número y tiene un hijo derecho y un hijo izquierdo.

Construimos el árbol de la siguiente forma:

Ponemos los primeros números en un cubo hasta llenarlo. Cuando tenemos $2b + 1$ números en el cubo reestructuramos el árbol. El cubo se vuelve nodo interno y llevará la mediana de los números del antiguo cubo. Además ponemos los números más pequeños que la mediana en un cubo a su izquierda (que estará entonces medio lleno) y los más grandes en otro cubo a su derecha. Por ejemplo, la permutación $(6, 3, 1, 4, 2, 5)$ con $c = 2$ tiene por árbol:



Coloreamos el árbol de la siguiente forma:

Tenemos $b+1$ colores. Tal como en el árbol m-ario vamos a colorear las brechas de las hojas.

Las brechas de una hoja que tiene $i + b - 1$ números (con $i = 1, ..., b + 1$) se colorean en color i.

Nótese que al contrario del árbol m-ario, aquí se colorean también las hojas llenas del árbol (las que tienen $2b$ números) y las hojas que siempre tienen por lo menos b números. El color i corresponde a un nodo hoja que tiene $i+b-1$ llaves insertadas, $i = 1, ..., b+1$. Una nueva llave que caiga en una hoja no rellena con $i+b-1$ llaves ($i \leq b$), aumenta su número de llaves a $i + b$. La correspondiente regla de crecimiento de la urna consiste en remplazar i bolas de color i por $i+1$ bolas de color $i+1$. La regla correspondiente a la división de un cubo es diferente: un nodo que se divide contiene $2b$ llaves con $2b+1$ brechas de color $b+1$. Si la inserción cae en una de estas brechas se produce la división en dos hojas cada una con b llaves es decir $b+1$ brechas de color 1, por lo tanto la urna gana $2b+2$ bolas de color 1.

Por ejemplo, nuestro árbol anterior tiene una urna correspondiente con tres bolas de color 2 y cuatro de color 1. La matriz asociada es:

$$A = \begin{pmatrix} -(b+1) & b+2 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & -(b+2) & b+3 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & -(b+3) & b+4 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & & & \ddots & & \vdots & \vdots \\ 0 & 0 & \dots & \dots & & & -2b & 2b+1 \\ 2b+2 & 0 & \dots & \dots & & & 0 & -(2b+1) \end{pmatrix}$$

Estudiando los valores propios de esta matriz obtenemos:

$$\begin{aligned} \chi_A(\lambda) &= det(A^T - \lambda I) \\ &= ((b+1)+\lambda)(-1)^b(b+2+\lambda)(b+3+\lambda)...(2b+\lambda)(2b+1+\lambda) \\ &+ (-1)^{b+1}(2b+2)(b+2)(b+3)...(2b)(2b+1) \end{aligned}$$

desarollando por la primera fila. Entonces

$$\chi_A(\lambda) = 0 \Leftrightarrow (\lambda + (b+1))((\lambda + (b+2))...(\lambda + (2b+1)) = \frac{(2b+2)!}{(b+1)!}.$$

Podemos observar que $\sum_{j=1}^{n} = a_{i,j} = 1 = \lambda_1$, $\forall i \in \{1, \dots, b+1\}$. Además si b es impar poniendo $\lambda = -3b - 3$ en el polínomio mínimo obtenemos

$$(-2b-2)(-2b-1)...(-b-2) = (-1)^{b+1}(2b+2)(2b+1)...(b+2) = \frac{(2b+2)!}{(b+1)!}$$

con lo cual $\lambda = -3b - 3$ es valor propio negativo.
Como en el caso del árbol m-ario de búsqueda 1 y eventualmente $-3b - 3$ son los únicos valores propios reales, como $\mathbb{C}$ es algebraicamente cerrado tenemos que todos los otros valores propios son complejos y vienen conjugados.

No desarrolaremos la prueba de que $Re(\lambda_2) < \frac{1}{2}Re(\lambda_1)$. Según Chern and Hwang (2001) esta desigualdad es cierta para $b \leq 58$, y con lo cual estamos en el caso de una urna ampliada si y solo si $b \leq 58$.

Ahora vamos a calcular el vector propio v asociado al valor propio principal $\lambda = 1$:

$$A^T v = v \Leftrightarrow \begin{cases} -(b+1)v_1 + (2b+2)v_{b+1} & = v_1 \\ (b+2)v_1 - (b+2)v_2 & = v_2 \\ \vdots & \vdots \\ (b+k)v_{k-1} - (b+k)v_k & = v_k \\ \vdots & \vdots \\ (2b+1)v_b - (2b+1)v_{b+1} & = v_{b+1} \end{cases}$$

de lo cual obtenemos que

$$v_i = \frac{b+2}{b+i+1} v_1.$$

La condición de normalizar el vector v se traduce en

$$v_1 + \frac{b+2}{b+3} v_1 + \frac{b+2}{b+4} v_1 + ... + \frac{b+2}{2b+2} v_1 = 1,$$

o sea

$$\begin{aligned} v_1 &= \frac{1}{1 + (b+2)(\frac{1}{b+3} + \frac{1}{b+4} + ... + \frac{1}{2b+2})} \\ &= \frac{1}{(b+2)(\frac{1}{b+2} + \frac{1}{b+3} + ... + \frac{1}{2b+2})} \\ &= \frac{1}{(b+2)(H_{2b+2} - H_{b+1})}, \end{aligned}$$

Entonces

$$v_i = \frac{1}{(b+i+1)(H_{2b+2} - H_{b+1})}$$

y utilizando el Teorema 3.5.1 tenemos:

$$\frac{X_n^i}{n} \to_P \frac{1}{(b+i+1)(H_{2b+2} - H_{b+1})}. \tag{4.1}$$

Ahora nos interesa conocer el comportamiento de S_n, el tamaño del árbol binario de páginas después de n pasos. Como en el caso del árbol m-ario de búsqueda denotamos I_n el número de nodos internos y L_n el número de hojas. Razonando como en el caso del árbol m-ario de búsqueda tenemos la relación siguiente (aqui $m = 2$):

$$I_n = L_n - 1,$$

entonces

$$S_n = I_n + L_n = 2L_n - 1 = 2 \sum_{k=1}^{b+1} \frac{X_n^{(k)}}{k+1}$$

y tomando límite usando 4.5 obtenemos

$$\frac{S_n}{n} \to_P 2 \sum_{k=1}^{b+1} \frac{1}{(b+k+1)(H_{2b+2} - H_{b+1})} \frac{1}{k+1}.$$

La normalidad asintótica del tamaño del árbol binario de páginas se establece en el siguiente resultado.

Teorema 4.4.1. *(Chern and Hwang (2001)). Sea S_n el tamaño del árbol binario de páginas ($c = 2b$, con $1 \leq b \leq 58$) después de n inserciones. Entonces,*

$$S_n^* = \frac{S_n - n/((2b+1)(H_{2b+1} - H_{b+1}))}{\sqrt{n}} \to_D N(0, \sigma_b^2).$$

La varianza asintótica σ_b^2 se puede determinar trás un cálculo tedioso que no desarrollaremos. Chern and Hwang (2001) prueban la normalidad asintótica para $1 \leq b \leq 58$ con el método de los momentos, también demuestran que para $b > 58$, S_n^* ya no tiene una distribución asintótica normal.

4.5. Árboles recursivos estándares

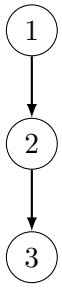
Los árboles recursivos tienen muchas aplicaciones para modelizar contagios, cadena de letras. Se aplica en filología, informática, etc... ,ver R.Smythe and H.Mahmoud (1996).

Esta teoría se ha desarollado en varias direciones pero ha aparecido recientemente que la teoría de las urnas de Pólya permitía unificar esos resultados.

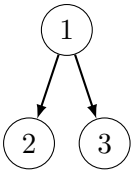
Vamos ahora a desarollar la teoría de los árboles recursivos estándares utilizando los modelos de urnas.

Empezamos el árbol con un nodo y en cada etapa se elige aleatoriamente de forma uniforme el padre del nodo entrante. El n-ésimo nodo entrante recibe el número n. Nótese que no hay restrición de números de hijos que pueda tener un nodo.

Por ejemplo, sólo hay dos árboles recursivos distintos de orden 3 que son los siguientes:



y



porque el nodo 2 se asociará necesariamente el 1 como padre, y el 3 se puede asignar el 1 (con probabilidad $\frac{1}{2}$) o el 2 (con probabilidad $\frac{1}{2}$).

Teorema 4.5.1. *(Najock and Heyde (1982)) Sea L_n el número de hojas (nodos que no tienen hijos) en un árbol recursivo estándar de tamaño n. Su distribución es:*

$$P(L_n = k) = \frac{1}{(n-1)!} E(k, n-1),$$

donde $E(k,n)$ es el número de Euler de primera especie (ver el apéndice) con $k = 1, ..., n-1$. Además

$$\frac{L_n - \frac{1}{2}n}{\sqrt{n}} \to N\left(0, \frac{1}{12}\right).$$

Demostración. Coloreemos las hojas del árbol en blanco y los otros nodos en azul. Cuando se elige un nodo blanco para ser el padre del nuevo entrante, la antigua hoja blanca se vuelve azul mientras que el nuevo entrante se colorea en blanco. Ahora si le toca a un azul recoger el nuevo entrante, el antiguo azul se queda azul y el nuevo entrante se colorea en blanco. En resumen en la urna correspondiente si cogemos una blanca añadimos una blanca y quitamos una, y añadimos una azul. Si cogemos una azul añadimos una blanca.

La matriz correspondiente es:

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

que es una matriz correspondiente a una urna de Friedmann.

Empezamos con una bola blanca en la urna ($W_0 = 1$, $B_0 = 0$). Si W_n es el número de bolas blancas después de n extracciones tenemos $L_n = W_{n-1}$.

Vamos a usar el Teorema 3.3.2 (el Teorema de Friedmann) con $s = 0$, $a = 1$, $\delta = \frac{\pi_0}{s-a} = \frac{1}{-1} = -1$, $\alpha = \frac{s+a}{s-a} = \frac{1}{-1} = -1$, y $\pi_n = n + 1$.

Se recuerda la ecuación que teníamos de la teoría de las urnas de Friedman, ver (3.2):

$$\chi_{n+1}(t) = \frac{e^{st}}{\pi_n} \left(1 - e^{-t(s-a)}\right)^{\alpha+1} \chi_n'(t)$$

que en nuestro caso es:

$$\chi_{n+1}(t) = \frac{\chi_n'(t)}{n+1}$$

y entonces por recurencia:

$$\chi_{n+1}(t) = \frac{1}{(n+1)!} \chi_0(t)^{(n)}$$

y ahora utilizando (3.1) es decir $\phi_n(t) = \left(1 - e^{-t(s-a)}\right)^{-\delta - \alpha n} \chi_n(t)$ con la condición inicial $\phi_0(t) = e^t$ obtenemos:

$$\phi_n(t) = \frac{1}{n!} \left(1 - e^t\right)^{n+1} \frac{d^n}{dt^n} \left(\frac{e^t}{1 - e^t}\right).$$

Ahora vamos a demostrar por recurrencia que

$$\frac{d^n}{dt^n}\left(\frac{e^t}{1-e^t}\right) = \frac{1}{(1-e^t)^{n+1}}\sum_{k=1}^{n} E(k,n)e^{kt}. \tag{4.2}$$

Para $n = 1$ tenemos

$$\frac{d^n}{dt^n}\left(\frac{e^t}{1-e^t}\right) = \frac{e^t}{(1-e^t)^2} = \frac{1}{(1-e^t)^2}e^t E(1,n)$$

porque solo hay una permutación de S_n que tiene un único ascenso.

Ahora suponemos que (4.2) se verifica para n, vamos a probarlo por $n+1$:

$$\begin{aligned}
&\frac{d^{n+1}}{dt^{n+1}}\left(\frac{e^t}{1-e^t}\right) = \frac{d}{dt}\left(\frac{1}{(1-e^t)^{n+1}}\sum_{k=1}^{n} E(k,n)e^{kt}\right)\\
&= \frac{1}{(1-e^t)^{n+2}}\sum_{k=1}^{n}\left(kE(k,n)\left(1-e^t\right)e^{kt} + (n+1)E(k,n)e^{(k+1)t}\right)\\
&= \frac{1}{(1-e^t)^{n+2}}\sum_{k=1}^{n}\left(kE(k,n)e^{kt} + (n+1-k)E(k,n)e^{(k+1)t}\right)\\
&= \frac{1}{(1-e^t)^{n+2}}\left(\sum_{k=1}^{n} kE(k,n)e^{kt} + \sum_{j=2}^{n+1}(n-j+2)E(j-1,n)e^{jt}\right)\\
&= \frac{1}{(1-e^t)^{n+2}}\left(E(1,n)e^t + \sum_{k=2}^{n}\left(kE(k,n) + (n-k+2)E(k-1,n)\right)e^{kt} + E(n,n)e^{(n+1)t}\right)\\
&= \frac{1}{(1-e^t)^{n+2}}\left(E(1,n+1)e^t + \sum_{k=2}^{n} E(k,n+1)e^{kt} + E(n+1,n+1)e^{(n+1)t}\right)
\end{aligned}$$

usando las propiedades de los números de Euler de primera especie que están desarrolladas en el apéndice se acaba la demostración de (4.2).

Según (4.2) tenemos

$$\phi_n(t) = \frac{1}{n!}\sum_{k=1}^{n} E(k,n)e^{kt}$$

y

$$\sum_{k=1}^{n-1} P(L_n = k)t^k = \sum_{k=1}^{n-1} P(W_{n-1} = k)t^k = \phi_{n-1}(ln(t)) = \frac{1}{(n-1)!}\sum_{k=1}^{n-1} E(k-1,n)t^k,$$

recordando que ϕ_n es por definición la función generatriz de W_n, se tiene

$$P(L_n = k) = \frac{1}{(n-1)!}E(k,n-1).$$

Luego utilizamos el Teorema 3.3.2 (Teorema de Friedmann) con $\rho = \dfrac{s-a}{s+a} = -1 < \dfrac{1}{2}$ y por lo cual

$$\frac{L_n - \frac{1}{2}n}{\sqrt{n}} \to N(0, \frac{1}{12}).$$

. □

Teorema 4.5.2. *(Gastwirth and Bhattacharya (1984)) Sea $S_{n,k}$ el tamaño del subárbol cuya raiz es la k-ésima inserción en un árbol recursivo de tamaño n. Cuando k y n crecen al infinito de tal forma que $\dfrac{k}{n} \to \rho$ tenemos*

$$S_{n,k} \to_D Geo(\rho).$$

Demostración. En la k-ésima etapa coloreemos el nuevo entrante, k, en blanco y todos los otros ($k-1$ nodos) en azul.

A cada nuevo entrante se le asigna el color de su padre (no se cambia de color ningún nodo). Es decir que la matriz asociada a este árbol es (después de la entrada del k-ésima nodo entrante):

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

que es un modelo de Pólya-Eggenberger con $s = 1$.

Las condiciones iniciales de la urna son: $W_0 = 1$, $B_0 = k-1$. Además $S_{k,n} = W_{n-k}$. Como anteriormente, denotamos W_n^* el número de bolas blancas extraidas de la urna en n sorteos obtenemos en nuestro caso la igualdad: $W_n = W_n^* + 1$.

Utilizando el Teorema 3.2.1 obtenemos:

$$\begin{aligned} P(W_{n-k}^* = j) &= \frac{j! * (k-1)k...(n-j-2)}{k(k+1)...(n-1)} \binom{n-k}{j} \\ &= \frac{(k-1)\Gamma(n-j-1)\Gamma(n-k+1)}{\Gamma(n)\Gamma(n-k-j+1)} \end{aligned}$$

Ahora utilizando la aproximación de Stirling:

$$\frac{\Gamma(x+r)}{\Gamma(x+s)} = x^{r-s} + o(x^{r-s-1})$$

obtenemos

$$\frac{\Gamma(n-k+1)}{\Gamma(n-k-j+1)} = (n-k)^j + o((n-k)^{j-1}) \text{ y } \frac{\Gamma(n-j+1)}{\Gamma(n)} = n^{-j-1} + o(n^{-j-2})$$

con lo cual

$$P(S_{k,n} - 1 = j) \approx \frac{k}{n^{j+1}}(n-k)^j \approx \frac{k}{n}\left(1 - \frac{k}{n}\right)^j$$

y por lo tanto

$$P(S_{k,n} = j) \to_{k,n\to\infty} (1-\rho)^{j-1}\rho.$$

□

Meir and Moon (1988) han estudiado distribuciones asociadas a grados de salida de un nodo (es decir cuantos hijos tiene un nodo, una hoja tiene grado de salida 0 por ejemplo). Gracias a los modelos de urnas podemos demostrar resultados sobre la distribución de los nodos de un cierto grado de salida.

Teorema 4.5.3. *(Janson (2005)) Sea $X_n^{(j)}$ el número de nodos de grado de salida j, $j = 0, ..., k$, en un árbol recursivo aleatorio de tamaño n y sea el vector $X_n = (X_n^{(1)}, ..., X_n^{(k)})^T$. Entonces,*

$$\frac{1}{\sqrt{n}}(X_n - n\mu_k) \to_D N_k(0_k, \Sigma_k)$$

donde $\mu_k = \left(\frac{1}{2}, \frac{1}{4}, ..., \frac{1}{2^{k+1}}\right)^T$, $0_k = (0, 0, ..., 0)^T$, y Σ_k una matriz de covarianzas.

Demostración. Vamos a aplicar el Teorema 3.5.2. Por eso coloreemos las hojas (los nodos de grado de salida 0) en color 1, a los nodos de grado de salida 1 le asociamos el color 2, y etc de tal forma que los nodos de grado de salida k se colorean de color $k+1$. Todos los nodos de grado de salida superior a k se colorean de color $k+2$.

Ahora si se elige un nodo de grado de salida j como padre del nuevo entrante ($j = 1, ..., k$), el padre ganará un hijo (tendrá entonces $j+1$ hijos). Lo cual equivale en la urna a perder una bola de color $j+1$ y ganar una de color $j+2$. Nótese que también ganamos una bola de color 1 porque el nuevo entrante será una hoja que colorearemos en color 1.

Si se elige un nodo de grado 0 como padre del nuevo entrante se pierde una bola de color 1 y se gana una de color 1 y otra de color 2.

El otro caso particular es si el nuevo entrante se va a asignar a un padre de grado de salida superior (estrictamente) a k, este nuevo hijo no cambia el color del padre pero añadimos una bola de color 1 (el hijo). Con lo cual tenemos:

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & \dots & 0 \\ 1 & -1 & 1 & 0 & \dots & 0 \\ 1 & 0 & -1 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & & \dots & 0 \\ & & & \ddots & & \\ 1 & 0 & 0 & \dots & 0 & 0 \end{pmatrix}$$

Se trata de la matriz de una urna ampliada. En efecto, tenemos que $\sum_{j=1}^{k+2} a_{i,j} = 1 = \lambda_1$, $\forall i \in \{1, \dots, k+2\}$.

Ahora vamos a comprobar la hipotesis $Re(\lambda_2) < \frac{1}{2}Re(\lambda_1)$.
Para encontrar los valores propios de A, estudiaremos su matriz transpuesta $A^T = B$. Miramos $\mathbb{C}^{k+2}$ como un subespacio de $l_1(\mathbb{N})$ y definimos $\pi_{k+2} : l_1(\mathbb{N}) \to \mathbb{C}^{k+2}$ la aplicación tal que

$$\pi_{k+1}(x_0, x_1, ...) = (x_0, ..., x_k, \sum_{i=k+1}^{\infty} x_i).$$

También definimos en $l_1(\mathbb{N})$,

$$S(x_0, x_1, ...) = (0, x_0, x_1, ...).$$

Denotaremos $u_1 = (1, 1, ..., 1)$ el vector propio asociado al valor propio principal $\lambda_1 = 1$. Por fin, sea $\delta_0 = (1, 0, ..., 0, 0)$

Tenemos entonces que

$$Bv = -v + (u_1.v)\delta_0 + \pi_{k+1}Sv, \forall v \in \mathbb{C}^{k+2}.$$

En efecto, nótese que $\pi_{k+1}Sv = \pi_{k+1}(0, v_1, v_2, ..., v_{k+2}) = (0, v_1, v_2, ..., v_k, v_{k+1} + v_{k+2})$.

Sea $E^{'} = \{v \in \mathbb{C}^{k+2} : u_1.v = 0\}$. En $E^{'}$ tenemos

$$B = -I + \pi_{k+1}S.$$

Entonces $\forall v \in E^{'}$, $(A + I)v = \pi_{k+1}S.v$ y observando que $\pi_{k+1}S\pi_{k+1} = \pi_{k+1}S$ obtenemos por recurrencia que:

$$(B + I)^k v = \pi_{k+1}S^k v, v \in E', k \geq 0.$$

En particular, $(B + I)^{k+1}v = \pi_{k+1}S^{k+1}v = 0, v \in E', k \geq 0$.

Entonces $B + I$ es nilpotente en E' y la restricción de B a E' tiene autovalor -1.
Como $E' + \mathbb{C}v_1 = \mathbb{C}^{k+2}$, tenemos que los valores propios de B son 1 e -1, -1 siendo de multiplicidad algebraica $k + 1$. Con lo cual se puede aplicar el Teorema 3.5.2.

Como siempre para aplicar el teorema 3.5.2 tenemos que calcular el vector propio de A^T asociado al valor propio principal 1 o sea:

$$A^T v = v \Leftrightarrow \begin{cases} v_2 + v_3 + ... + v_{k+2} & = v_1 \\ v_1 - v_2 & = v_2 \\ v_2 - v_3 & = v_3 \\ \vdots & \vdots \\ v_i - v_{i+1} & = v_{i+1} \\ \vdots & \vdots \\ v_{k+1} & = v_{k+2} \end{cases}$$

del cual obtenemos que $\forall i \in \{1, \ldots, k+1\}$,

$$v_i = \left(\frac{1}{2}\right)^{i-1} v_1$$

y $v_{k+2} = \left(\frac{1}{2}\right)^{k} v_1$.

La condición de normalizar el vector v se traduce como:

$$v_1 + \frac{1}{2}v_1 + \frac{1}{4}v_1 + ... + \left(\frac{1}{2}\right)^{k} v_1 + \left(\frac{1}{2}\right)^{k} v_1 = 1$$

o sea

$$v_1 \left(\sum_{i=0}^{k} \left(\frac{1}{2}\right)^{i} + \left(\frac{1}{2}\right)^{k} \right) = 1,$$

de lo cual deducimos que $v_1 = \frac{1}{2}$ entonces

$$v_i = \left(\frac{1}{2}\right)^{i} \text{ si } i \in \{1, \ldots, k+1\} \text{ y } v_{k+2} = \left(\frac{1}{2}\right)^{k+1}$$

Entonces utilizando el Teorema de 3.5.1 tenemos:

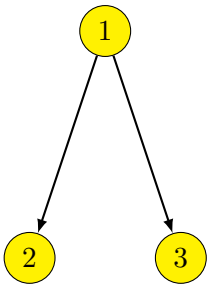
$$\frac{X_n^{(i)}}{n} \to_P \left(\frac{1}{2}\right)^{i}$$

y finalmente utilizando el Teorema 3.5.2 acabamos la demostración.

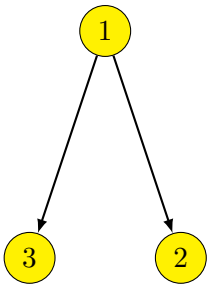
□

4.6. Árbol recursivo estándar orientado

En el caso del árbol recursivo estándar, no se tomaba en cuenta la orientación del árbol. Por ejemplo, los árboles

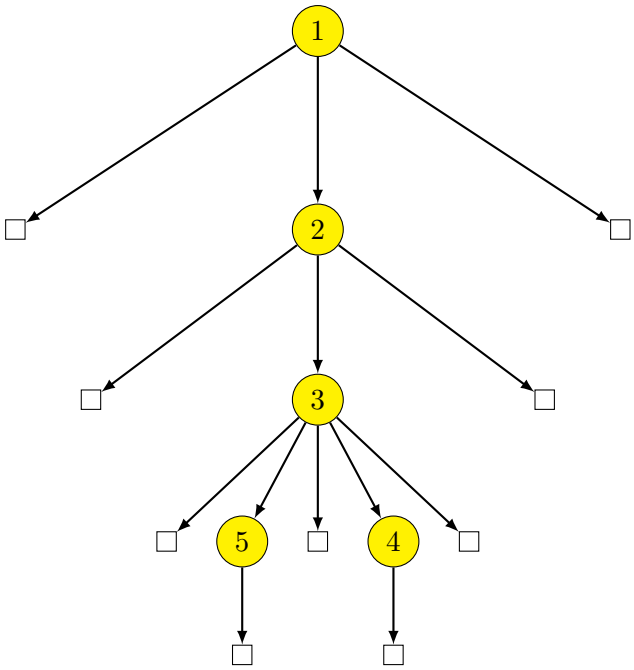


y



solo son dos formas distintas de dibujar el mismo árbol recursivo estándar. En efecto, hemos de recordar cómo se construye el árbol recursivo estándar, en los dos casos el uno ha sido elegido como padre del dos y del tres. Pero nuestro método de construción del árbol no daba restricción sobre donde se dibuja el nuevo entrante.

Ahora consideremos otra forma de construir el árbol que tome en cuenta la orientación. Para eso ya no será el padre del nuevo entrante que estará elegido al azar de manera uniforme sino la brecha (tal que la habíamos definido en la sección del árbol m-ario de búsqueda) donde se insertará el nuevo entrante. Por ejemplo el siguiente árbol es un árbol recursivo estándar orientado y extendido (las brechas son los cuadrados).



Se puede encontrar en Mahmoud H. and J. (1993) la distribución exacta y asíntotica de las hojas de tal árbol. Nosotros demostraremos solamente el siguiente resultado.

Teorema 4.6.1. *Mahmoud H. and J. (1993) En un árbol recursivo estándar orientado de tamaño n, el número de hojas (denotado L_n) verifica*

$$\frac{L_n - \frac{2}{3}n}{\sqrt{n}} \to N\left(0, \frac{1}{9}\right).$$

Demostración. Aquí vamos a colorear las brechas porque son los objetos que se eligen al azar de manera uniforme cuando se construye el árbol recursivo orientado (y no los nodos como era el caso en el árbol recusivo estándar como explicado antes). Coloreamos en blanco las brechas que provienen de una hoja, y en azul las otras brechas. Entonces si el nuevo entrante se inserta en una brecha blanca (o sea cogemos una bola blanca en la urna) perdemos una brecha blanca y ganamos otra, además aparecen dos brechas azules a la izquierda y a la derecha de la antigua brecha. Ahora, si el nuevo entrante se inserta en una brecha de color azul, perdemos una brecha de color azul y ganamos una de color blanca y dos de color azul a la izquierda e a la derecha del nuevo nodo. Con lo cual el esquema de la urna asociada es

$$\begin{pmatrix} 0 & 2 \\ 1 & 1 \end{pmatrix}$$

lo cual es un esquema de una matriz de Bagchi y Pal. Entonces aplicamos el Teorema 3.4.1, con $a - c = -1 < 1 = K$ de aquí el resultado (notando que $W_n = L_n$) □

5
Conclusión

En este trabajo solo hemos podido desarrollar una parte muy limitada de las aplicaciones que tienen las urnas de Pólya. Tampoco hemos estudiado en toda su generalidad el concepto de urna de Pólya ya que no hemos desarrollado resultados con entradas aleatorias. Acabaremos este trabajo enseñando un ejemplo de applicación de las urnas de Pólya y de los árboles aleatorios a la química. Este problema esta propuesto como ejercicio en el libro de Mahmoud (2009).

El árbol recursivo no uniforme con grado de salida acotado ha sido introducido para modelizar el crecimiento de un componente químico. Suponemos que un polímero crece de forma similar a un árbol recursivo. Un nodo atrae la próxima molecula según su peso o afinidad que son determinados por su grado de salida. Un nodo de grado i tiene una afinidad proporcional a $4 - i$. Los nodos de grado 4 son saturados, entonces una nueva molécula sólo se puede insertar en un nodo que tiene ya grado 1,2 o 3. En este modelo los nodos con grados pequeño son "hambrientos" y tienen más probabilidad de atraer a una nueva molecula. ¿Cuál es la proporción asíntotica de moleculas "hambrientas" (las hojas) en el polímero (el árbol) ?.

Ya no se trata de un árbol recursivo uniforme, porque en su construción habíamos añadido la regla de que algunos nodos tienen más probabilidad que otros de atraer al nuevo entrante. Con lo cuál no vamos a colorear los nodos sino las brechas que son los objectos uniformes. En efecto, coloreamos los nodos externos procedente de un nodo de grado de salida 0 (o sea los nodos externos hijos de las hojas) de color blanco, los nodos externos procedentes de nodos de grado de salida 1 se colorean en azul, los nodos externos procedentes de nodos de grado de salida 2 se colorean en verde y los nodos externos procedentes de nodos de grado de salida 3 se colorean en rojo.

De aquí obtenemos que la matrix asociada a esta urna es

$$A = \begin{pmatrix} 0 & 3 & 0 & 0 \\ 4 & -3 & 2 & 0 \\ 4 & 0 & -2 & 1 \\ 4 & 0 & 0 & -1 \end{pmatrix}$$

lo cual es una matriz ampliada à la Smythe. En efecto trás calculo del polinomio mínimo obtenemos:

$$Sp(A) = \{3, -2, -3, -4\}$$

donde $\lambda_1 = 3$ es valor propio principal (nótese que A es de suma constante sobre las filas). Se verifica la hipotesis $\lambda_2 = -2 < \frac{1}{2}\lambda_1 = \frac{3}{2}$.

Entonces solo nos queda por determinar el vector propio izquierdo normalizado asociado a λ_1.

Resolviendo $A^T v = 3v$ obtenemos $v = \left(2v_2, v_2, \frac{2}{5}v_2, \frac{1}{10}v_2\right)$ y normalizando obtenemos $v_2 = \frac{10}{35}$.

Aplicando el Teorema 3.5.1 obtenemos

$$\frac{W_n}{n} = \frac{L_n}{n} \to_P 3 \times \frac{20}{35} = \frac{60}{35}.$$

6
Apéndice

6.1. Números de Euler de primera especie

Los números de Euler de primera especie, denotado $E(k,n)$ cuenta el número de permutaciones de $\{1,2,...,n\}$ que tienen exactamente k rachas ascendentes (una racha de longitud 1 se considera ascendente).

Por ejemplo, la permutación $(1,2,3)$ tiene una racha ascendente mientras que la permutación $(1,3,2)$ tiene dos.

Tenemos la relación:

$$E(k,n) = kE(k,n-1) + (n-k+1)E(k-1,n-1)$$

En efecto, insertamos n en una permutación de $\{1,...,n-1\}$ en cualquiera de las n brechas posibles. Al realizar esta inserción observamos que el número de ascensos o bien permanece constante o bien aumenta en una unidad, tal como veremos a continuación. Si una permutación de $\{1,2,...,n-1\}$ tiene k ascensos y la inserción de n se realiza en una brecha correspondiente al final de un ascenso, entonces el número de ascensos de la nueva permutación es también k. Obsérvese que esto puede hacerse de k formas.

Si una permutación de $\{1,2,...,n-1\}$ tiene $k-1$ rachas ascendentes entonces la inserción de n en una brecha que no sea el final de una racha ascendente o al final hace que el número de rachas se incremente en una unidad, obsérvese que esto puede hacerse de $n-k+1$ formas.

6.2. Martingalas

Una sucesión de variables aleatorias es una martingala si y solo si se verifica:

1) $\forall n \geq 1,\ E\left[|X_n|\right] < \infty$

2) $E\left[X_n \mid X_{n-1}\right] = X_{n-1}$

Vamos ahora a presentar el teorema central límite para las martingalas, necesitamos dos condiciones sobre $\nabla X_n = X_n - X_{n-1}$.

1) La condición de Lindeberg: $\forall \epsilon > 0$,

$$\sum_{k=1}^{n} E\left[(\nabla X_k)^2 1_{\{|\nabla X_k| > \epsilon\}} \mid \mathcal{F}_{k-1}\right] \to_P 0.$$

2) Una condición de varianza $Z-$condicional es:

$$\sum_{k=1}^{n} E\left[(\nabla X_k)^2 \mid \mathcal{F}_{k-1}\right] \to_P Z.$$

El teorema central del límite demostrado en Hall and Heyde (1980) viene dado por:

Teorema 6.2.1. *Sea X_n una martingala que satisface la condición de Linderberg y una condición de varianza $Z-$condicional entonces $\frac{X_n}{\sqrt{n}}$ converge en distribución a una variable aleatoria cuya función de distribución es $E\left[exp(-Zt^2)\right]$.*

Bibliografía

A. Bagchi and A. Pal. Asymptotic normality in the generalized Pólya-Eggenberger urn model with applications to computer data structures. *SIAM Journal on Algebraic and Discrete Methods*, (6):394–405, 1985.

S. Bernstein. Sur un problème du schéme des urnes à composition variables. *C.R Dokl. Acad. Sci. URSS*, (28):5–7, 1940.

H. Chern and H. Hwang. Phase change in random m-ary search trees and generalized quicksort. *Random Structures and Algorithms*, (19):316–358, 2001.

L. Devroye. Limit laws for local counters in random binary search trees. random structures and algorithms. (2):303–316, 1991.

F. Eggenberger and G. Pólya. Über die Statistik Verketetter Vorgäge. *Zeitschrift f¨ur Angewandte Mathematik und Mechanik.*, (1):279–289, 1923.

D. Freedman. Bernard friedman's urn. *Annals of Mathematical Statistics*, (36):956–970, 1965.

B. Friedman. A simple urn model. *Communications of Pure and Applied Mathematics*, (2):59–70, 1949.

J. Gastwirth and P. Bhattacharya. Two probability models of pyramids or chain letter schemes demonstrating that their promotional claims are unreliable. *Operations Research*, (32):527–536, 1984.

P. Hall and C. Heyde. *Martingale Limit Theory and Its Applications.* Academic Press, N.Y., 1980.

H.Mahmoud. Evolution on random search trees. *Wiley, New-York*, 1992.

Robert T. Smythe, Hosam M. Mahmoud. Probabilistic analysis of bucket recursive trees. *Theoretical Computer Science*, (144):240–241, 1995.

S. Janson. Asymptotic degree distribution in random recursive trees. *Random Structures and Algorithms*, (26):69–83, 2005.

H. Mahmoud. On rotations in fringe-balanced binary trees. *Information Processing Letters*, (65):41–46, 1998.

Hosam M. Mahmoud. *Pólya Urn Models*. Chapman & Hall/CRC, The Georges Washington University, District of Columbia, USA, 2009.

Smythe R. Mahmoud H. and Szymański J. On the structure of plane-oriented recursive trees and their branches. *Random structure and algorithm*, (4):151–176, 1993.

A Meir and J. Moon. Recursive trees with no nodes of out-degree one. *Congressus Numerantium*, (66):49–62, 1988.

D. Najock and C. Heyde. On the number of terminal vertices in certain random trees with an application to stemma construction in philology. *Journal of Applied Probability*, (19):675–680, 1982.

A. Panholzer and H. Prodinger. An analytic approach for the analysis of rotations in fringe-balanced binary search trees. *Annals of combinatorics*, (2):173–184, 1998.

R.Smythe and H.Mahmoud. A survey of recursive trees. *Theory of Probability and Mathematical Statistics*, (51):1–29, 1996.

V. Savkevich. Sur le schéma des urnes à composition variables. *C.R Dokl. Acad. Sci. URSS*, (28):8–12, 1940.

R. Smythe. Central limit theorems for urn models. *Stochastic Processes and Their Applications.*, (65):115–137, 1996.

R.P Stanley. *Enumerative Combinatorics, Vol 2*. Cambridge Studies in Advanced Mathematics, 1999.